

GenCode version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2006, 11:16:24 ; Search time 80 Seconds
(without alignments)
2784.563 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAAPPPAPRRARRRRERED.....RMKLAAPLESTPLKXRGTA 507

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21:.*
1: geneseqp19808:.*
2: geneseqp19808:.*
3: geneseqp20008:.*
4: geneseqp20018:.*
5: geneseqp20028:.*
6: geneseqp20038:.*
7: geneseqp20038:.*
8: geneseqp20048:.*
9: geneseqp20058:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2670	100.0	507	ADJ96642	Adj96642 Human Nim
2	2499	93.6	499	ADV97865	Adv97865 Murine pr
3	2409	90.2	462	AAB65654	Aab65654 Novel pro
4	2409	90.2	462	AD129261	Ad129261 Human MAR
5	1889	69.6	357	AAW78691	Aaw78691 Human pro
6	1580	59.2	326	AAB66605	Aab66605 Human h15
7	1534.5	57.5	535	AAB65657	Aab65657 Novel pro
8	1534.5	57.5	535	AAB33483	Aab33483 Human pro
9	1534.5	57.5	535	ABBO5696	Abbo5696 Human bra
10	1534.5	57.5	535	ADD47447	Add47447 Human pro
11	1534.5	57.5	535	ADG10732	Adg10732 Human STR
12	1534.5	57.5	535	AD129264	Ad129264 Human MAR
13	1534.5	57.5	535	ADOS8659	Ados8659 Human reg
14	1534.5	57.5	535	AEAE1758	Aeae1758 Human nuc
15	1534.5	57.5	581	AAU87231	Aau87231 Human cen
16	1534.5	57.5	581	AD154546	Ad154546 Novel hum
17	1529.5	57.3	535	AAAB92901	Aaab92901 Human pro
18	1529.5	57.3	535	AAAG7420	Aaag7420 Amino aci
19	1508.5	56.5	535	AAW77229	Aaw77229 Amino aci
20	1444	54.1	501	AAAB4157	Aaab4157 Human pan
21	1360.5	51.0	281	AAAB5655	Aaab5655 Novel pro
22	1360.5	51.0	281	AD129262	Ad129262 Mouse MAR
23	1346	50.4	258	AEAE1759	Aeae1759 Human nuc
24	1310	49.1	434	ADG10844	Adg10844 Human STR

25	1174.5	44.0	637	4	ABR57826	Abbr57826 Drosophil
26	1170	43.8	293	6	ABP96552	Abp96552 Human pro
27	1106.5	41.4	378	4	AAB65658	Aab65658 Novel pro
28	1106.5	41.4	378	8	AD129265	Ad129265 Mouse MAR
29	951.5	35.6	308	4	AAB36408	Aab36408 Secretd
30	946	35.4	340	4	AAU17233	Aau17233 Novel sig
31	946	35.4	340	4	AAU87535	Aau87535 Novel cen
32	946	35.4	340	7	ADB93941	Adb93941 Human nov
33	946	35.4	340	8	AD154850	Ad154850 Novel hum
34	882.5	33.1	752	7	ADC87261	Adc87261 Human GPC
35	756	28.3	302	4	AAU17257	Aau17257 Novel sig
36	756	28.3	302	4	AAU87232	Aau87232 Novel cen
37	756	28.3	302	7	ADB93965	Adb93965 Human nov
38	756	28.3	302	8	AD154547	Ad154547 Novel hum
39	336	14.8	2091	8	ADN61449	Adn61449 Human KPP
40	335	14.8	779	5	ABB97147	Abb97147 Human tum
41	335	14.8	1345	5	AAE25097	Aae25097 Human kin
42	335	14.8	2193	6	ABR42219	Abrr42219 Human pro
43	395	14.8	2217	9	ADX98203	Adx98203 lysine de
44	395	14.8	2219	8	ADN00364	Adn00364 Novel hum
45	395	14.8	2245	8	ADJ96651	Adj96651 Human Nim

ALIGNMENTS

RESULT 1
ADJ96642
ID ADJ96642 standard; protein; 507 AA.
XX
AC ADJ96642;
XX
DT 06-MAY-2004 (first entry)
XX
DB Human Nim-A related protein kinase NRBP2 protein SeqID 99.
XX
KW kinase; human; tyrosine protein kinase; serine/threonine protein kinase;
XX PTK; STK; gene therapy; cancer; immune-related disease;
XX cardiovascular disease; brain; neuronal associated disease; metabolic;
XX inflammatory disorder; cytosolic; neuroprotective; immunomodulator;
XX antiinflammatory; enzyme; Nim-A related protein kinase; NRBP2.
OS Homo sapiens.
OS 57.
XX
XX WO2004006838-A2.
XX PN
XX 22-JAN-2004.
XX PD
XX PF 15-JUL-2003; 2003WO-US021730.
XX PR 15-JUL-2002; 2002US-0395632P.
XX XX
PA (SUGB-) SUGEN INC.
XX
PI Whyte D, Manning G, Caenepeel S;
XX WPI; 2004-122753/12.
XX DR N-PSDB; ADJ96576.
XX
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
XX Claim 1; SEQ ID NO 99; 366pp; English.
XX
XX This invention relates to a novel isolated, enriched or purified nucleic
XX acid molecule that encodes a kinase polypeptide. Specifically, it relates
XX to human tyrosine and serine/threonine protein kinases (PTK's and STK's),
XX as well as protein kinase-like enzymes. The present invention describes
XX screening methods to identify agonists, antagonists and antibodies that
XX can be used to modulate the activity or function of the mammalian kinase
XX enzymes. As such, these compositions can be used for gene therapy

CC purposes to treat diseases or disorders including cancer, immune-related
 CC diseases, cardiovascular disease, brain or neuronal associated disease,
 CC metabolic and inflammatory disorders. Accordingly, they exhibit
 CC cytosolic, neuroprotective, immunomodulator and anti-inflammatory
 CC activities. This polypeptide sequence is a human kinase protein sequence
 CC of the invention.

XX Sequence 507 AA;

Query Match 100.0%; Score 2670; DB 8; Length 507;

Best Local Similarity 100.0%; Pred. No. 1,2e-250; Mismatches 0; Indels 0; Gaps 0;

Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPEPAPRRAREREREDESEDESDESDILESPPCGRMQRREOVNQGMPGLQSTFLAMD 60
 DB 1 MAPEPAPRRAREREREDESEDESDESDILESPPCGRMQRREOVNQGMPGLQSTFLAMD 60
 QY 61 TEEGVEVWVWNEHFGDRKAPAAHEEKTQVFEQVLVDHPNIVKLHKWLDTSEACAVI 120
 DB 61 TEEGVEVWVWNEHFGDRKAPAAHEEKTQVFEQVLVDHPNIVKLHKWLDTSEACAVI 120
 QY 121 FITEVSSGSLKQFLKTKKNNKAMNARAKMCTQIISALSPFNACSPPIIHGULTSDT 180
 DB 121 FITEVSSGSLKQFLKTKKNNKAMNARAKMCTQIISALSPFNACSPPIIHGULTSDT 180
 QY 181 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 240
 DB 181 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 240
 QY 241 TAVVIDFSFGMCALMAVLEIQNGDTRVTEBAIARARHSLSDPMNREFILCCLARDPAR 300
 DB 241 TAVVIDFSFGMCALMAVLEIQNGDTRVTEBAIARARHSLSDPMNREFILCCLARDPAR 300
 QY 301 PSASHLLFHRVLFVYHSLKTLAAHCFIOHQYLMPEVVEKTKAMDHLAVLAELPRPRP 360
 DB 301 PSASHLLFHRVLFVYHSLKTLAAHCFIOHQYLMPEVVEKTKAMDHLAVLAELPRPRP 360
 QY 361 PLQWRYSEVSEFMELEKFLIEDVANGIYPLMNPATRPGLPRVLADPPEEVOAKPTTDEP 420
 DB 361 PLQWRYSEVSEFMELEKFLIEDVANGIYPLMNPATRPGLPRVLADPPEEVOAKPTTDEP 420
 QY 421 PFSERKRYIQMOCNERSEDKARWHLTLVLVEDRLHQLTYDLPDTSADQLASELVHY 480
 DB 421 PFSERKRYIQMOCNERSEDKARWHLTLVLVEDRLHQLTYDLPDTSADQLASELVHY 480
 QY 481 GFLHEDDRMKLAAPLESTFLKYRGTOA 507
 DB 481 GFLHEDDRMKLAAPLESTFLKYRGTOA 507

RESULT 2
 ADV97865
 ID ADV97865 standard; protein; 499 AA.

AC ADV97865;

DT 24-MAR-2005 (first entry)

DE Murine protein kinase enzyme Seq 185.

XX enzyme; protein kinase modulator; bioinformatics; diagnostic; cancer;
 XX central nervous system disease; psychiatric disorder;
 XX neurological disorder; neurodegenerative disease; metabolic disorder;
 XX cardiovascular disease; inflammatory disorder; gene therapy; cytostatic;
 XX antiinflammatory; analgesic; endocrine-gen.; nootropic; tranquilizer;
 XX hypotensive; hypertensive; neuroprotective; antiparkinsonian; vitruide;
 XX fungicide; antibacterial; antidiabetic; anorectic; antidiabetic; antineoplastic;
 XX ophthalmological; antiinflammatory; antiallergic; antineoplastic;
 XX antiaesthetic; osteopathic; antipsoriatic; immunosuppressive;
 XX cardiovascular-gen.; vasotropic; antiallergic; gastrointestinal-gen.;
 XX cns-gen.
 XX Mus sp.

XX WO200500200-A2.
 PN 06-JAN-2005.
 XX 07-MAY-2004; 2004WO-US014421.
 XX 09-MAY-2003; 2003US-0469014P.
 XX (SUGEN) SUGEN INC.
 PA Caenepeel S, Manning G, Charyczak G, Grigoriev I;
 PI WPI; 2005-066508/07.
 DR N-PSDB; ADV97751.
 PT New isolated, enriched, or purified kinase nucleic acids and
 PT polypeptides, useful for diagnosing or treating, e.g. cancers,
 PT neurological and neurodegenerative diseases, cardiovascular disease, or
 PT inflammatory disorders.

PS Claim 3; SEQ ID NO 185; 300bp; English.

CC This invention relates to novel isolated, enriched or purified nucleic
 CC acid molecules that encode kinase polypeptides. Specifically, it refers
 CC to a bioinformatics strategy used to identify mammalian members of the
 CC protein and lipid kinase families. The present invention provides methods
 CC for identifying a substance that modulates the activity of a kinase
 CC polypeptide, as well as a method for the detection of a kinase
 CC acid in a sample as a diagnostic tool for a disease or disorder.
 CC Furthermore, it describes generation of a knock-out mouse whose genome is
 CC disrupted by recombination at a nucleic acid sequence such that it
 CC produces a phenotype, relative to the wild-type, that exhibits an absence
 CC of kinase activity. The nucleic acids and polypeptides given in the
 CC specification are useful for the diagnosis and treatment of cancer,
 CC central or peripheral nervous system diseases, psychiatric and neurological
 CC disorders, neurodegenerative diseases, metabolic disorders,
 CC cardiovascular disease or inflammatory disorders. As such, they can be
 CC used for gene therapy purposes and compositions exhibit cytostatic,
 CC antiinflammatory, analgesic, endocrine-gen., nootropic, tranquilizer,
 CC hypotensive, hypertensive, neuroprotective, antiparkinsonian, vitruide,
 CC fungicide, antibacterial, antidiabetic, anorectic, antidiabetic, antineoplastic,
 CC ophthalmological, antiinflammatory, antiallergic, antineoplastic,
 CC antiaesthetic, osteopathic, antipsoriatic, immunosuppressive, cardiac-
 CC gen, vasotropic, antiallergic and gastrointestinal-gen. This polypeptide
 CC is a murine protein kinase of the invention.

XX Sequence 499 AA;

Query Match 93.6%; Score 2499; DB 9; Length 499;

Best Local Similarity 93.9%; Pred. No. 5.5e-234; Mismatches 12; Indels 8; Gaps 2;

Matches 476; Conservative 11; Mismatches 12; Indels 8; Gaps 2;

QY 1 MAPEPAPRRAREREREDESEDESDESDILESPPCGRMQRREOVNQGMPGLQSTFLAMD 60
 DB 1 MAPEPAPRRAREREREDESEDESDESDILESPPCGRMQRREOVNQGMPGLQSTFLAMD 58
 QY 61 TEEGVEVWVWNEHFGDRKAPAAHEEKTQVFEQVLVDHPNIVKLHKWLDTSEACAVI 120
 DB 61 TEEGVEVWVWNEHFGDRKAPAAHEEKTQVFEQVLVDHPNIVKLHKWLDTSEACAVI 118
 QY 121 FITEVSSGSLKQFLKTKKNNKAMNARAKMCTQIISALSPFNACSPPIIHGULTSDT 180
 DB 121 FITEVSSGSLKQFLKTKKNNKAMNARAKMCTQIISALSPFNACSPPIIHGULTSDT 178
 QY 181 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 240
 DB 181 IFIQHNGLIKIGSVWHRIFSNALRPPTALPDDLSPIRAEREELRNLFPPPEYGEVADG 232
 QY 241 TAVVIDFSFGMCALMAVLEIQNGDTRVTEBAIARARHSLSDPMNREFILCCLARDPAR 300
 DB 241 TAVVIDFSFGMCALMAVLEIQNGDTRVTEBAIARARHSLSDPMNREFILCCLARDPAR 292

PT New antisense compound targeted to a nucleic acid molecule encoding
PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
PT expression of MARK3 or for treating cancer or Alzheimer's disease.

XX Disclosure, SEQ ID NO 181; 233bp; English.

XX The invention relates to a compound comprising a sequence comprising 8-80
CC base pairs (bp) targeted to a nucleic acid encoding MARK3
CC (MAP/microtubule affinity-regulating kinase 3), that specifically
CC hybridises with the nucleic acid encoding MARK3 and inhibits expression
CC of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
CC composition comprising the compound and a carrier or diluent, inhibiting
CC the expression of MARK3 in cells or tissues, treating an animal having or
CC suspected of having a disease or condition associated with MARK3 and
CC screening for an antisense compound. The antisense oligonucleotide is
CC useful for preparing a composition for treating hyperproliferative
CC disorder, particularly cancer and neurodegenerative diseases e.g.
CC Alzheimer's disease. The present sequence is a MARK3 associated protein
CC included in the figures but not mentioned anywhere else in the
CC specification.

XX Sequence 462 AA;

Query Match 90.2%; Score 2409; DB 8; Length 462;
Best Local Similarity 98.3%; Pred. No. 2,9e-225;
Matches 460; Conservative 2; Mismatches 0; Indels 6; Gaps 1;

QY 40 RREOVNQGMPGLOSTFLAMDTEEGEVVWNLHFGDRAPAFANEKIQTVESQVLVDH 99
DB 1 QREKVNQGMPPGLOSTFLAMDTEEGEVVWNLHFGDRAPAFANEKIQTVESQVLVDH 60
QY 100 PNIIVLHKYMLDTESEACAVIFITEYVSSGSLKQFLKTKKNHKKANAAAMKRMCTQILS 159
DB 61 PNIIVLHKYMLDTESEACAVIFITEYVSSGSLKQFLKTKKNHKKANAAAMKRMCTQILS 120
QY 160 AISFLHACSPPIIHGNLISDTFIQHNGLIKIGSVWHRIEVSALRPPTALPDLSPIRA 219
DB 121 AISFLHACSPPIIHGNLISDTFIQHNGLIKIGSVWHRIEVSALRPPTALPDLSPIRA 174
QY 220 EBEELRNHFFPEYGEVADGTAVIDIFSGMCALENAVLEIQNDGTVTEEAIRARHS 279
DB 175 EBEELRNHFFPEYGEVADGTAVIDIFSGMCALENAVLEIQNDGTVTEEAIRARHS 234
QY 280 LSDPNMREFILCCLARDPARPSAHSILFHRVLFVHSLKLAACFIQHOYLMENYVE 339
DB 225 LSDPNMREFILCCLARDPARPSAHSILFHRVLFVHSLKLAACFIQHOYLMENYVE 294
QY 340 EKTAMDLHAVALAEILPRPRRPLQWRYSSEVSFMEIDKFLFEDVRNGIYPLMNPAAATRPICL 399
DB 295 EKTAMDLHAVALAEILPRPRRPLQWRYSSEVSFMEIDKFLFEDVRNGIYPLMNPAAATRPICL 354
QY 400 PRVLAAPPEEVOKATPTPEPDSERKVIQOQCNLERSDQARHNTLLVLERLRHQ 459
DB 365 PRVLAAPPEEVOKATPTPEPDSERKVIQOQCNLERSDQARHNTLLVLERLRHQ 414
QY 460 LTYDDLPTDSADLASLVHGFLEHDDRMKLAFLFESTFLKYRGTOA 507
DB 415 LTYDDLPTDSADLASLVHGFLEHDDRMKLAFLFESTFLKYRGTOA 462

RESULT 5
AAM78691
ID AAM78691 standard; protein; 357 AA.

XX AAM78691;

XX 06-NOV-2001 (first entry)

DE Human protein SEQ ID NO 1353.

KM Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukaemia;

KM nervous system disorder; arthritis; inflammation.

XX Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US004098.

XX 03-FEB-2000; 2000US-00469914.

XX 27-APR-2000; 2000US-00560875.

XX 20-JUN-2000; 2000US-00598075.

XX 19-JUL-2000; 2000US-00620325.

XX 01-SEP-2000; 2000US-00654936.

XX 15-SEP-2000; 2000US-00663361.

XX 20-OCT-2000; 2000US-00693325.

XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Dirmacac RT, Abundi V, Zhou P, Xu C, Cao Y,
XX Ma Y, Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW,
XX Xue AJ, Yang Y, Wejhtman T, Goodrich R,
XX WPI; 2001-476283/51.

XX N-PSDB; AAK51824.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
XX in diagnosis and gene therapy.

XX Claim 20; Page 3598-3599; 6221dp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
XX encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activity/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
XX (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication

XX Sequence 357 AA;

Query Match 69.6%; Score 1858; DB 4; Length 357;
Best Local Similarity 97.8%; Pred. No. 9.4e-172;
Matches 355; Conservative 1; Mismatches 1; Indels 6; Gaps 1;

QY 145 MNABAMKWCQIILSALSFHACSPPIIHGNLISDTFIQHNGLIKIGSVWHRIEVSALR 204

DB 1 MNABAMKWCQIILSALSFHACSPPIIHGNLISDTFIQHNGLIKIGSVWHRIEVSALR 57

QY 205 PPTLAPDDLRSPPIAERBELNLHFFPEYGEVADGTAVIDIFSGMCALENAVLEIQNG 264

DB 58 ---ALPDLRSPPIAERBELNLHFFPEYGEVADGTAVIDIFSGMCALENAVLEIQNG 114

QY 265 DTRVTEEAIRARHSISDPNMRREFILCCLARDPARPSAHSILFHRVLFVHSLKLAH 324

DB 115 DTRVTEEAIRARHSISDPNMRREFILCCLARDPARPSAHSILFHRVLFVHSLKLAH 174

QY 325 CFIOHOYLMENYVEEKTAMDLHAVALAEILPRPRRPLQWRYSSEVSFMEIDKFLFEDVRNG 384

DB 175 CFIOHOYLMENYVEEKTAMDLHAVALAEILPRPRRPLQWRYSSEVSFMEIDKFLFEDVRNG 234

QY 385 IYPLMNPAAATRPICLPRVLAAPPEEVOKATPTPEPDSERKVIQOQCNLERSDQARW 444

DB 235 IYPLMNPAAATRPICLPRVLAAPPEEVOKATPTPEPDSERKVIQOQCNLERSDQARW 294

QY 445 HTLTLVLVEDRLHRLQTLTYDLLPTDSAQDLASLVLHYGFLHEDDRMKLAFLSTFLKXRG 504
DB 295 HTLTLVLVEDRLHRLQTLTYDLLPTDSAQDLASLVLHYGFLHEDDRMKLAFLSTFLKXRG 354
QY 505 TQA 507
DB 355 TQA 357

RESULT 6

AAB66605
ID AAB66605 standard; protein; 326 AA.

AC AAB66605;

DT 04-APR-2001 (first entry)

DE Human h15993 protein.

KM Human; protein kinase; cell growth; tumour; cancer; immune; inflammatory;
respiratory; haematological; bone disorder.

OS Homo sapiens.

PN WO200100879-A1.

PD 04-JAN-2001.

PF 30-JUN-2000; 2000WO-US018291.

PR 30-JUN-1999; 99US-00345473.

PR 01-MAY-2000; 2000US-00562480.

XX (MILL-) MILLENNIUM PHARM INC.

PI Hodge MR, Meyers R, Williamson M;

DR WPI; 2001-061977/07.

PT New protein kinase polypeptides, nucleic acids and anti-kinase
antibodies, useful for diagnosing and treating e.g. cancer, inflammatory,
immune, cardiovascular and bone disorders.

PS Claim 1; Fig 13; 93pp; English.

CC The present invention relates to human protein kinase. The proteins are
CC from new human genes termed h12832, h14138, h14633, h15990, h15993,
CC h16341 and h2252. The proteins may be used to identify modulators of
CC their activity. The proteins may also be used to derive products for the
CC treatment of cellular growth related disorders, malignancies, cancers,
CC immune, inflammatory, respiratory, haematological and bone-related
CC disorders

SQ Sequence 326 AA;

Query Match 59.2%; Score 1580; DB 4; Length 326;
Best Local Similarity 97.4%; Pred. No. 1e-144;
Matches 301; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 MAAPAPAPRAREREREREDESESDIEESPGRMOKRREOVNQGMMPGLOSTFLAMD 60
DB 12 MAAPAPAPRAREREREREDESESDIEESPGRMOKRREOVNQGMMPGLOSTFLAMD 71
QY 61 TEEGVEVVMNELHFGDRKAPAAHEEKIQTFPEQLVLDHPNIVKJLHKYWLDTSEACARYI 120
DB 72 TEEGVEVVMNELHFGDRKAPAAHEEKIQTFPEQLVLDHPNIVKJLHKYWLDTSEACARYI 131
QY 121 FITEVYSSSLKQPLKTKKQNKAMNARAKWKCTQIISLSTLAACSPPIIHGNTSPT 180
DB 132 FITEVYSSSLKQPLKTKKQNKAMNARAKWKCTQIISLSTLAACSPPIIHGNTSPT 191
QY 181 FIGHNGLIKIGSVMHRIISNLRPTALPDRLRSPIRAREBELRLVHFFPEYGEVADG 240

DB 192 FIGHNGLIKIGSVMHRIISN-----ALPDRLRSPIRAREBELRLVHFFPEYGEVADG 245
QY 241 TAVDIFSEGMCALENVALEIQTNQDTRVTEBAIARARHSLDPNNKEFLCCLARDPAR 300
DB 246 TAVDIFSEGMCALENVALEIQTNQDTRVTEBAIARARHSLDPNNKEFLCCLARDPAR 305
QY 301 PSAHSLTFH 309
DB 306 PSVHSLTFH 314

RESULT 7

AAB65657
ID AAB65657 standard; protein; 535 AA.

AC AAB65657;

DT 27-MAR-2001 (first entry)

DE Novel protein kinase, SEQ ID NO: 184.

KM Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
immunosuppressive; cardiac; renal; antiinflammatory; antiashtmatic;
dermatological; antidiabetic; antifertility; gene therapy; vaccine;
immune disorder; cardiovascular disease; neurodegenerative disease;
cancer; autoimmune disorder; stroke; inflammatory bowel disease;
inflammatory pelvic disease; multiple sclerosis; psoriasis.

OS Homo sapiens.

PN WO200073469-A2.

PD 07-DEC-2000.

PF 26-MAY-2000; 2000WO-US014842.

PR 26-MAY-1999; 99US-0136503P.

XX (SUGB-) SUGEN INC.

PI Plowman GD, Martinez R, Whyte D, Sudersanam S;

DR WPI; 2001-032161/04.
N-PSDB; AAR44684.

PT Nucleic acids encoding kinase polypeptides, useful for diagnosing and
treating immune-related diseases and disorders, cardiovascular disease,
neurodegenerative diseases and/or cancers.

PS Claim 10; Fig 1; 310pp; English.

CC The present sequence is a novel protein kinase. The novel protein kinases
CC and the nucleic acids that encode them may be used in the treatment and
CC diagnosis of diseases associated with inappropriate kinase expression
CC such as immune-related diseases and disorders, cardiovascular disease,
CC neurodegenerative diseases and/or cancers. The nucleic acids and
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays. The kinase polypeptides may be used as antigens in the production
CC of antibodies of kinase expression and activity. Anti-kinase antibodies
CC and kinase antagonists may also be used to down regulate kinase
CC expression and activity. Diseases related to kinase expression and
CC activity include rheumatoid arthritis, atherosclerosis, autoimmune
CC disorders, complications of organ transplantation, myocardial infarction,
CC immune disorders, cardiomyopathies, strokes, renal failure, oxidative-
CC stress related disorders, chronic inflammatory bowel disease, chronic
CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,
CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive
CC disorders

SQ Sequence 535 AA;

Query Match 57.5%; Score 1534.5; DB 4; Length 535;
Best Local Similarity 61.6%; Pred. No. 5.6e-140;

[illegible][illegible]

AB05696
 ID AB05696 standard; protein; 535 AA.
 AC AB05696;
 DT 30-APR-2002 (first entry)
 DE Human brain derived protein clone fbr2_76d18.
 KW Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
 KW gene therapy.
 OS Homo sapiens.
 PN WO200198454-A2.
 PD 27-DEC-2001.
 PF 25-APR-2001; 2001WO-IB02050.
 PR 25-APR-2000; 2000US-0199380P.
 XX (GEHU-) GERMAN HUMAN GENOME PROJECT.
 PA Wiemann S;
 PI MPI, 2002-055860/07.
 DR N-PSDB; ABA93734.
 XX Human cDNA sequences and clones derived from human fetal brain, fetal
 PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
 PT screening and therapy.
 XX Claim 1; Page 216-217; 611pp; English.
 PS
 XX The present invention describes assemblies and computer readable media
 CC comprising novel human cDNA sequences and clones derived from human
 CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
 CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
 CC present invention which encode the proteins given in AB05662 to
 CC AB05729. The human cDNA sequences and clones can be used in gene
 CC therapy. The clones may be used in a variety of applications, for example
 CC they may be used in profiling assays, for providing large arrays of human
 CC genetic material for implementing large-scale screening strategies and
 CC for treating diseases via gene therapy procedures
 CC
 CC Sequence 535 AA;
 SQ
 Query Match 57.5%; Score 1534.5; DB 5; Length 535;
 Best Local Similarity 61.6%; Pred. No. 5,6e-140;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

DB 330 VPSILKLAHNCIVGHQHMIPENALBEITKQMTSAVLAEIPAGRGREPVQTIYSQSPALE 389
 QY 374 LDKFLEDVNGIYPLMNFATRPGLPRLVADPPPEV-----QKAKTPPEPPDSETRK 427
 DB 390 LDKFLEDVNGIYPLTAF-----GLPRPQOQOQEEVTSPPVPSVKTPTPPAVETRK 443
 QY 428 VIQMCNLEERSDKRMHTLLVLVEDRLHROLTYDLPTDSAQDLASLYHYGFLEHD 487
 DB 444 VVLMQCNLESVEBGVGHLLTLKKEDKLRHLSCLMNMENIPELAELVOLGFISEAD 503
 QY 488 RMKLAAFLESTFLKY 502
 DB 504 QSRUTSULEETLANKF 518
 RESULT 10
 ADD47447
 ID ADD47447 standard; protein; 535 AA.
 AC ADD47447;
 DT 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)
 XX Human Protein NP_037524, SEQ ID NO 13142.
 DE Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNJ; Chung.
 XX Homo sapiens.
 OS Unidentified.
 PN WO2003016475-A2.
 PD 27-FEB-2003.
 PF 14-AUG-2002; 2002WO-US025765.
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEHU-) GEN HOSPITAL CORP.
 PA (PARB) BAYER AG.
 PI Woolf C, D'urso D, Befort K, Costigan M;
 DR MPI, 2003-268312/26.
 DR GENBANK; NP_037524.
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 PS Example 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the

Qy 374 LKFLFEDVANGIYPLNNFATRLPLGLPVLAAPPPEV-----QKAKTPPEPDSSETRK 427
 Db 390 LKFLFEDVANGIYPLTAFA-----GLPRPOQPOQEEVTSVPVPSVKTPTPEPAVEYTRK 443
 Qy 428 VIOMQCNLERSDEKARWHLTLVLVEDRLHROLTYDLLPTDSAQDLASLHVHGFLEDD 487
 Db 444 VVLMQCNISVEGVGHHTLLKLEDKLNRLHSLCDLMENENIPBLAELVQLGFISEAD 503
 Qy 488 RMKLAFLSESTFLKY 502
 Db 504 QSRLTSLSEETLNKF 518
 RESULT 12
 ADI29264
 ID ADI29264 standard; protein; 535 AA.
 AC ADI29264;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Human MARK3-associated protein #47.
 XX
 KW Human; antisense gene therapy; MARK3;
 KW MAP/microtubule affinity-regulating kinase 3; cancer;
 KW Alzheimer's disease; neurodegenerative disorder;
 KW hyperproliferative disorder; cytostatic.
 XX
 OS Homo sapiens.
 XX
 PN US2003232771-A1.
 XX
 PD 18-DEC-2003.
 XX
 PF 17-JUN-2002; 2002US-00174319.
 XX
 PR 17-JUN-2002; 2002US-00174319.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Ward DT, Freiler SM, Dobie KW,
 DR WPI: 2004-052188/05.
 DR N-PSDB; ADI29382.
 XX
 PT New antisense compound targeted to a nucleic acid molecule encoding
 PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
 PT expression of MARK3 or for treating cancer or Alzheimer's disease.
 XX
 PS Disclosure; SEQ ID NO 184; 233bp; English.
 XX
 CC The invention relates to a compound comprising a sequence comprising 8-80
 CC base pairs (bp) targeted to a nucleic acid encoding MARK3
 CC (MAP/microtubule affinity-regulating kinase 3), that specifically
 CC hybridizes with the nucleic acid encoding MARK3 and inhibits expression
 CC of MARK3, i.e. is an antisense oligonucleotide (AO). Also included are a
 CC composition comprising the compound and a carrier or diluent, inhibiting
 CC the expression of MARK3 in cells or tissues, treating an animal having or
 CC suspected of having a disease or condition associated with MARK3 and
 CC screening for an antisense compound. The antisense oligonucleotide is
 CC useful for preparing a composition for treating hyperproliferative
 CC disorder, particularly cancer and neurodegenerative diseases e.g.
 CC Alzheimer's disease. The present sequence is a MARK3 associated protein
 CC included in the figures but not mentioned anywhere else in the
 CC specification.
 CC
 XX
 SQ Sequence 535 AA;
 Query Match 57.5%; Score 1534.5; DB 8; Length 535;
 Best Local Similarity 61.6%; Pred. No. 5.6e-140;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;
 Qy 17 BREDSSEDSILBESPCGRWQKRRQVNVQGNMPLGLSTFLAMDTEGVEVVMNELHFGD 76

Db 44 EEESEDSSEILBESPCGRWQKRRQVNVQGNMPLGLSTFLAMDTEGVEVVMNELHFGD 103
 Qy 77 RKAFAAHEEKIOTVEQOLVVDHPNIVKLAKTWLDTSEACARVIFITEYVSSGSLKQFLK 136
 Db 104 RKQYKIQEERKAVAFPNQILQELHINIVKFKHWADIKENKARVIFITEYVSSGSLKQFLK 163
 Qy 137 KTKKNNKANNARAKMRCCTQIISALSFACSPRIIHGMLTSDTIFIQHNGLIKIGSVH 196
 Db 164 KTKKNNKNNERAKMRCCTQIISALSYLHSCDPIIHGMLTCDTIFIQHNGLIKIGSV-- 221
 Qy 197 RIFSNALRPPTALPDLLRSPIRAREELNLHFPPEYGEVAD-GTAVDIFSGMCALRM 255
 Db 222 -----APDTINNHVKTCKREQKILFPABEYSEVNVTTAVDITFSGMCALRM 269
 Qy 256 AVLEIQTNGDTR-VTEBAIARARHSISDNMEFFILCCLARDPARPSASHLLFHVLEP 314
 Db 270 AVLEIQNGESSYVPEAIISSAIIQLLEDPLOREFFIQCLQSEPARRPTARELLFHPALPE 329
 Qy 315 VHSIKLLAHCFTIQOYLMPENVVEKTKAMDHLAVLALPR-PRRPLQMKYSEVSFME 373
 Db 330 VPSLKLAAHCTVGHQHMIPENALBEITTKMDTSAYLABIPAGPREPVQTLYSQSPALPE 389
 Qy 374 LKFLFEDVANGIYPLNNFATRLPLGLPVLAAPPPEV-----QKAKTPPEPDSSETRK 427
 Db 390 LKFLFEDVANGIYPLTAFA-----GLPRPOQPOQEEVTSVPVPSVKTPTPEPAVEYTRK 443
 Qy 428 VIOMQCNLERSDEKARWHLTLVLVEDRLHROLTYDLLPTDSAQDLASLHVHGFLEDD 487
 Db 444 VVLMQCNISVEGVGHHTLLKLEDKLNRLHSLCDLMENENIPBLAELVQLGFISEAD 503
 Qy 488 RMKLAFLSESTFLKY 502
 Db 504 QSRLTSLSEETLNKF 518
 RESULT 13
 ADOS8699
 ID ADOS8699 standard; protein; 535 AA.
 AC ADOS8699;
 XX
 DT 15-JUL-2004 (first entry)
 XX
 DE Human regulatory molecule HRM-20.
 XX
 KW cytostatic; immunomodulator; agonist; antagonist; gene therapy;
 KW human regulatory molecule; HRM; disease development; cell proliferation;
 KW immune response; cancer.
 XX
 OS Homo sapiens.
 XX
 PN US2002058264-A1.
 XX
 PD 16-MAY-2002.
 XX
 PF 26-SEP-2001; 2001US-00840787.
 XX
 PR 23-SEP-1997; 97US-00933750.
 PR 20-JAN-1999; 99US-00234613.
 PR 03-MAR-2000; 2000US-00518865.
 XX
 PA (INCY-) INCYTE PHARM INC.
 XX
 PI Lal P, Hallman JL, Bandman O, Shah P, Au-Young J, Yue H;
 PI Guegler KJ, Corley NC;
 DR WPI: 2004-459763/43.
 DR N-PSDB; ADOS8748.
 XX
 PT New human regulatory molecules, useful in the diagnosis and treatment of
 PT cancer and immune disorders.
 XX

PS Claim 1; SEQ ID NO 20; 116bp; English.

XX The invention describes human regulatory molecules (HRM) (I) selected
 CC from a group comprising the fully defined amino acid sequences of SEQ ID
 CC NOs: 1-49. Also described are: an isolated polynucleotide (II) comprising
 CC a nucleic acid sequence encoding (i) or the complement of the
 CC polynucleotide (SEQ ID NOs:50-98); a composition comprising (II) and a
 CC reporter molecule; an expression vector containing (II); a host cell
 CC containing the vector; detecting (M1) expression of a nucleic acid in a
 CC sample; screening (M2) a plurality of molecules to identify a ligand;
 CC diagnosing (M3) a disease associated with gene expression in a sample
 CC containing nucleic acids; a composition comprising (I) and a
 CC pharmaceutical carrier or a labeling moiety; screening (M4) a plurality
 CC of molecules to identify a ligand; preparation and purification of
 CC antibodies; an antibody which specifically binds to (i); and detecting
 CC protein expression in a sample. The new human regulatory protein
 CC molecules which are expressed during disease development and the
 CC polynucleotides which encode them satisfies a need in the art by
 CC providing compositions which are useful in the diagnosis and treatment of
 CC diseases associated with cell proliferation, particularly immune
 CC responses and cancers. This is the amino acid sequence of a human
 CC regulatory molecule.

XX Sequence 535 AA;

Query Match 57.5%; Score 1534.5; DB 8; Length 535;
 Best Local Similarity 61.6%; Pred. No. 5.6e-140;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

QY 17 EREDSSESDILSESPCGRMQKRREOVNQNMFGLOSTFLAMDTREGVYVWNLHFGD 76
 DB 44 EEESESESEILSESPCGRMQKRREOVNQNMFGLOSTFLAMDTREGVYVWNLHFGD 103
 QY 77 RKAFAHEKIQTFEBOVLVDHPNIVYKHKWLDTSACAVIFITEVSSGSLKQFLK 136
 DB 104 RKNYKLGEBKVAEVDNLQLEHLNIVYKHKWLDTSACAVIFITEVSSGSLKQFLK 163
 QY 137 KTKKHKAMNARAKMKTCTQILSALSFHACSPPIIHGNLTSDTIFIOHNGLIKIGSVH 196
 DB 164 KTKKHKMTNEMKMKWCTQILSALSYLHSCDPIIHGNLTSDTIFIOHNGLIKIGSV-- 221
 QY 197 RIFSNALRPPTALPDOLSPRIAREERELNHHFPEVEGVAD-GTAVDIFSGMCALEM 255
 DB 222 -----APDTINNHVKTCEEQKULHFPAEYGEVNTVTAVDIYSFGMCALEM 269
 QY 256 AVLEIOTNGDTR-VYBEATARAARHSDPMREFFLCCARPARPRAHSLHFRVLF 314
 DB 270 AVLEIOTNGESSYVPEALISSAIQLLEDPLOREFFLOKQSQSPARPRTRRELLFHPALFE 329
 QY 315 VHSKLTLAAHCFIOHQYLMPEVNEVEKTKAMDHLAVIAPR-PRRPLQMRYSSEVFW 373
 DB 330 VPSLKLTLAAHCTVGHQHMIPENALBEITQNMOTSAVLAIPAGPQREPIQTIQSQSPAL 389
 QY 374 LDKFLEADVNGIYPLMNPATRPGLPRVLAPRPEEV-----QKATPTPPEPDSERK 427
 DB 390 LDKFLEADVNGIYPLAF-----GLPRPOQEQEVEVTPVVPSPKPTPPAPAEVETRK 443
 QY 428 VVOMCNTERSEBDKARWHLTLVLVEDRLHROLTVDILPTDQAOLASVHYGFLHED 487
 DB 444 VVLMQCNTERSEVGGVGHRLTLKLKEDKLNRLSCDLMENENIPELAELVOLGFISEAD 503
 QY 488 RMLKALFESTFLAKY 502
 DB 504 QSRITSLEETLNKF 518

RESULT 14
 ID AEA61758 standard; protein; 535 AA.
 AC AEA61758;
 XX
 DT 11-AUG-2005 (first entry)

XX Human nuclear receptor binding protein (NRBP).

DE Nuclear receptor binding protein; GTPase modulator; cancer; neoplasm;
 KM cytosolic; gene therapy; antibody therapy; drug screening; diagnosis.
 XX Homo sapiens.
 OS WO2005052130-A2.
 PN 09-JUN-2005.
 XX 23-NOV-2004; 2004WO-US039708.
 PF 24-NOV-2003; 2003US-0524541P.
 PR (EXEL-) EXELIXIS INC.
 PA Kadyk LC, Francis GR, Lickteig K;
 PI MPI, 2005-417982/42.
 DR N-PSDB; AEA61756.
 DR REFSEQ; NP_037524.

Identifying a candidate RAC pathway modulating agent using a nuclear
 receptor binding protein (NRBP) polypeptide or nucleic acid, useful for
 diagnosing or treating cancer.

Example 2; SEQ ID NO 3; 51bp; English.

XX The present sequence is the protein sequence of human nuclear receptor
 CC binding protein (NRBP), a ubiquitously expressed, multidomain adapter
 CC protein. The inventors have discovered genes that modify the RAC GTPase
 CC pathway in *Caenorhabditis elegans* and identified their human orthologs as
 CC NRBP. The invention provides methods for using these RAC modifier genes
 CC and polypeptides to identify NRBP-modulating agents that are candidate
 CC therapeutic agents that can be used in the treatment of disorders
 CC associated with defective or impaired RAC function and/or NRBP function.
 CC Preferred NRBP-modulating agents specifically bind to NRBP polypeptides
 CC and restore RAC function. Other preferred NRBP-modulating agents are
 CC nucleic acid modulators such as antisense oligomers and RNAi that repress
 CC NRBP gene expression or product activity. NRBP modulating agents may be
 CC evaluated by *in vitro* or *in vivo* assay for molecular interaction with an
 CC NRBP polypeptide or nucleic acid. Agents that produce a change in the
 CC activity of these assay system relative to controls are identified as
 CC candidate RAC modulating agents. The assay system may be cell-based or
 CC cell-free. NRBP modulating agents include NRBP-related proteins (e.g.
 CC dominant negative mutants and biotherapeutics), NRBP-specific antibodies,
 CC NRBP-specific antisense oligomers and other nucleic acid modulators, and
 CC chemical agents that specifically bind to or interact with NRBP or
 CC compete with an NRBP binding partner. The screening assay may be an
 CC apoptosis assay, a cell proliferation assay, an angiogenesis assay or a
 CC hypoxic induction assay. A method for diagnosing a disease, especially
 CC cancer, in a patient uses a probe for NRBP expression.

SEQ Sequence 535 AA;

Query Match 57.5%; Score 1534.5; DB 9; Length 535;
 Best Local Similarity 61.6%; Pred. No. 5.6e-140;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

QY 17 EREDSSESDILSESPCGRMQKRREOVNQNMFGLOSTFLAMDTREGVYVWNLHFGD 76
 DB 44 EEESESESEILSESPCGRMQKRREOVNQNMFGLOSTFLAMDTREGVYVWNLHFGD 103
 QY 77 RKAFAHEKIQTFEBOVLVDHPNIVYKHKWLDTSACAVIFITEVSSGSLKQFLK 136
 DB 104 RKNYKLGEBKVAEVDNLQLEHLNIVYKHKWLDTSACAVIFITEVSSGSLKQFLK 163
 QY 137 KTKKHKAMNARAKMKTCTQILSALSFHACSPPIIHGNLTSDTIFIOHNGLIKIGSVH 196
 DB 164 KTKKHKMTNEMKMKWCTQILSALSYLHSCDPIIHGNLTSDTIFIOHNGLIKIGSV-- 221

DB:	6	Gaps:	1
US-10-618-941-99 (1-507) x AR448352 (1-4048)			
QY	1	MeCaLaAaLProGluProAlaProAlaGTAaGAlaArgGluuArgGluuArgGluuArg	20
Db	282	ATGGCGCGCCCGGAGCGCGCGCGGAGCGCGCGGAGCGGAGCGGAGCGGAGCGGAGCG	341
QY	21	GlusSerGluAProGluSerAerPilleuGluGluSerProCysGlyArgTrpGluuArg	40
Db	342	GAGAGCGAGGAGCGAGAGCGACATCTCGAGAGAAAGCCCGTGGTGCTGGCAAAAGCGA	401
QY	41	ArgGluGluValAserGlnGlyAserMetProGlyLeuGlnSerThrPheLeuAmetAap	60
Db	402	CGGAGCGAGGTAAACCAAGAGGAAACAATGCCAGGCGCTTCAGAGCACTTCTCAAGCAATGAGAC	461
QY	61	ThrgGluGluGlyValGluValVal1TPAserGluLeuHisPheGlyAerAerGlyValAphe	80
Db	462	ACGAGAGAGGGGGGTAGAGGTGGTGTGTAACAGACTCACTTCGAGACAGAGAGGCTTTC	521
QY	81	AlaAlaHisGluGluGluGlySerGlnThrValPheGluGluLeuValLeuValAerHisApro	100
Db	522	CGCGCGCAGAGGAGAGAGATCCAGACCGTGTTGAGAGAGCTGGTGCTGGTGGACACCCG	581
QY	101	AerHisLeuValGlyLeuHisGlySerTrpPleuAerThrSerGluAcyValAerGlyAlle	120
Db	582	AACATCGTGAAGTTGGACAAGTACTCGCTGGAACTCTGAAGCGCTGGCGGAGGTCAATC	641
QY	121	PheHisLeuGluGluTrpValSerSerGlySerLeuGlyGlnPheLeuGlySerThrGlyVal	140
Db	642	TTTCATCAACAGATGACTGTGTCATAGGCGAGCTTAAGCAATCTCTCAAAAAGACCAAGAG	701
QY	141	AerHisGlyValAerAerAerAa1aArgAla1aTrpGlySerTrpCysThrGlnHisLeuSerAla	160
Db	702	AACCAACAAGCCATGAAAGCGCGCGCTGGAAACGCTGGTGACACGAGATCTGTCTGGCG	761
QY	161	LeuSerPheLeuHisValAcySerSerProPro1leHisGlyValLeuThrSerAerPThr	180
Db	762	CTCAGAGTTCCTGACCGCTGACCGACCGCCCAATATCAACGGAACTGACACCGACACACC	821
QY	181	HisPheHisGlnHisAerGlnGlyLeuHisGlySerVal1aTrpHisAerGluPheSer	200
Db	822	ATCTTCATTCACAGACACAACGGCTCATACAGATGGGCTCGTGTTGGACCGAATCTTCTCC	881
QY	201	AerHisAlaLeuArgProProThrAlaLeuProAerAerPleuAerSerPro1leArgAlaGlu	220
Db	882	AAT-----GCACTTCCAGATGATCTCCGAAGCCCAATCCGCGGTGAG	923
QY	221	ArgGluGluLeuArgAerLeuHisAerPheProProGluTrpGlyGluValAlaAerGly	240
Db	924	CGAGAGGAATCTCGGAACCTGCACTTCTTCCCGCCAGAGATGAGAGGTGGCGGATGGGG	983
QY	241	ThrAlaValAerPlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeuGluuile	260
Db	984	ACCGCTGTGACACTTCTCTCTTGAGATGTGTGCGGTGAGATGGCTGTACTGTGAATTC	104
QY	261	GlnThrAerGlnYAsrPThrArgVal1ThrGluGluAla1leHisAerGluAerHisSerLeu	280
Db	1044	CAGACCAAAAGGGGACACCGCGGTCAACAGAGAGAGGCAATCTGCGCCGACGACTGGCTG	110
QY	281	SerAerProAerMetAerGluPhe1leLeuCysAlaLeuAlaArgAerProAlaAerGarg	300
Db	1104	AGTGACCCCAACAATCGCGGAGTTCAATCTTGGCTGGCTGGCCCGGAGCCTTGGCCGCGG	116
QY	301	ProSerAlaHisSerLeuAerPheHisAerGluValLeuPheGluValHisSerLeuGlyLeu	320
Db	1164	CCCTGTGCCCAACGCTCTCTTCCACCGGTGCTCTTCAAGGTGCACTGCTGAAGCTC	122
QY	321	LeuAlaAlaHisGlyAerPheHisGlnHisGlnTrpLeuMetProGluuAerVal1aGluGlu	340
Db	1224	CTGGCAGGCCACATGCTTCAATCCAGACACAGTACTACTCAATGCTGAGATGTGTGGAGAG	128
QY	341	LyThrLyValAerAerAerLeuHisValAlaValLeuAlaGluLeuAerProArgAerGargPro	360

Db	1284	AAAGCAGAGCCATGAGACTTCAAGCGGCTTTGGGCGAGCTTCCCGGACCCCGCAGGCCC	1344
Qy	361	ProLeuGlnTrpArgTyrSerGluValSerPheMetGluLeuAspLysPheLeuGluAsp	380
Db	1344	CGCGTGCAGTGGCGGGTACTCGGAAGTCTCTTACATGAGCTGGACAAATTCCTGGAGGAT	1403
Qy	381	ValArgbAngIyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGluPro	400
Db	1404	GTCAGGAATGGAATCTACCCACTGATGAACTTTGGACGACTCGACCCCTGGGCTGGCC	1463
Qy	401	ArgValIleuAlaProProGluGluValGlnLysAlaLysThr-ProThrProGluPro	420
Db	1464	CGTGGCTGGGCCCCACCCCGGAGAGGTCCAAAGGCGAAGACCCCGGACGACAGGC	1522
Qy	420	OPheAspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAs	440
Db	1524	CTTTGACTCTGAGACCGAAGAAAGTCTATCCAAATGCAAGTGCACCTGGAGAGAGGAGA	1583
Qy	440	PLysAlaArgTrpHisLeuThrLeuLeuValLeuGluAspArgLeuHisArgGlnLe	460
Db	1584	CAAGCGCGGCTGGCATCTCACTGCTCTTCTGGTCTGGAGAACCGGCTGCACCGGACAGCT	1644
Qy	460	uThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisTyr	480
Db	1644	GACCTACGACCTGCTCCCAACGAGCAGGCGCCAGAGACTCGCTCGAGGCTGGCACCTA	1703
Qy	480	rgLysPheLeuHisGluAspArgMetLysLeuAlaAlaPheLeuGluSerThrPheLe	500
Db	1704	TGGCTCTCTCCACAGAGACGACCGGATGAAGCTGGCGCCCTTCTCGAGAGAGCACTTCT	1763
Qy	500	uLysTyrArgGlyThrGlnAla 507	
Db	1764	CAAGTACCGTGGAGACCCAGGCC 1785	
RESULT 2			
LOCUS	BC071605	3892 bp	mRNA linear PRI 25-JUN-2004
DEFINITION	Homo sapiens hypothetical protein LOC340371, mRNA (cDNA clone		
	IMAGE:4375917), partial cds.		
ACCESSION	BC071605		
VERSION	BC071605.1	GI:48735259	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo		
REFERENCE	1 (bases 1 to 3892)		
AUTHORS	Straussberg, R.L., Felngold, E.A., Grouse, L.H., Derge, J.G.,		
	Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,		
	Altschul, S.F., Zeeberg, B., Bucow, K.H., Schaefer, C.F., Bhat, N.K.,		
	Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,		
	Ditschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,		
	Stepien, O.M., Soares, M.B., Bonaldi, M.F., Casavant, T.L.,		
	Scheetz, T.E., Brownstein, M.J., Udell, T.B., Tothbielki, S.,		
	Carrinini, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,		
	Abraham, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,		
	McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,		
	Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W.,		
	Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,		
	Fateh, J., Hellton, E., Kettelman, M., Madan, A.C., Rodriguez, S.,		
	Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,		
	Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,		
	Dickson, M.C., Rodriguez, A.C., Gilmwood, J., Schnut, J., Myers, R.M.,		
	Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalun, D.E.,		
	Schuerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.		
	Generation and initial analysis of more than 15,000 full-length		
	human and mouse cDNA sequences		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
PUBMED	12479932		
REFERENCE	2 (bases 1 to 3892)		
AUTHORS	Straussberg, R.		

QY 439 GIUAAPLYAAlaArgTRPHisLeuThriLeuLeuValLeuGIUAAPArgLeuHisArg 458
DB 1600 GAGGACAGAGCCGCGCTGGACATCTCACTCTGCTTCTGCTGGAAGACCGGCTGGACCG 1659
QY 459 GInLeuThrTyraSPLeuLeuProThraSPSerAlaGInaSPLeuAlaSerGIuLeuVal 478
DB 1660 CAGCTGACCTTACGACTGCTCCCAACGAGCGGCCAGAGACTGCGCTCGAGGCTCGTG 1719
QY 479 HIRTYRGlyPheLeuHisLeuLeuAPArgMetIySLeuAlaAlaPheLeuGIuSerThr 498
DB 1720 CACTATGGCTTCTCCACAGAGACGACCGGATGAAGCTGCGCTTCTCGAGAGACACC 1779
QY 499 PheLeuIyETyraRGlyThriGlnAla 507
DB 1780 TTCCTCAAGTACCGTGGACCCAGGCC 1806

RESULT 3
LOCUS CQ729503 1548 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15437 from Patent WO02068579.
ACCESSION CQ729503
VERSION CQ729503.1 GI:42300828
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kite, such as nucleic acid arrays, comprising a majority of
humaneons or transcripts, for detecting expression and other uses
thereof
Patent: WO 02068579-A 15437 06-SEP-2002;
JOURNAL PE Corporation (NY) (US)
FEATURES
source
1. 1548
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 4,286-206 Length: 1548
Score: 2517.50 Matched: 489
Percent Similarity: 94.22% Conservative: 0
Best Local Similarity: 94.22% Mismatches: 1
Query Match: 94.29% Indels: 29
DB: 6 Gaps: 4

US-10-618-941-99 (1-507) x CQ729503 (1-1548)

QY 12 ATGgiUAArgGIUAArgGIUAArgGIUAArgGIUAArgGIUAArgGIUAArgGIUAArg 31
DB 7 CGGGAACGGGAACGGGAACGGGAACGGGAACGGGAACGGGAACGGGAACGGGAACGGGA 66
QY 32 SerProCysGlyAlaGlyTrpGlnIySarGATArgGInValaSPGInGlyAAsnMetProGly 51
DB 67 ACCCGCTGTGCTGCTGCGCAAAAGCGACCGGAGCAGGTAAACCAAGGGAACATGCGCAGGG 126
QY 52 LeuGIuSerThrPheLeuAlaMetAspThrGluGluGlyValaGIuValValaTrpAsnGlu 71
DB 127 CTTTCAGAGCACTTCTTACCCATGACACGAGAGAGAGGGGTAGAGGTGTGTGGAACGAG 186
QY 72 LeuHisPheGlyAAspArgIySaAlaPheAlaAlaHisLeuGIuIyLeuIleGInThrValPhe 91
DB 187 CTCACCTTGGAGACAGAAAGGCTTCGCGCGCGACGAGAGAAATCCAGACCGTGTTC 246
QY 92 GIuGIuLeuValaLeuValaSPHisProAsnIleValaIyLeuHisIyTyTrpLeuAsp 111
DB 247 GAGGAGCTGTGTGTGTGTGACCAACCGGACATCGTAAGTTGCACAAATGACTGCTGAT 306

QY 112 ThrSerGIuAlaCysAlaArgValaIlePheIleThriGluIyValaSerSerGlySerLeu 131
DB 307 ACCTTGAGGCTCTGGCGAGGGGTATCTTCAATCAGAGGTACGTCTCATCAGGACGCTC 366
QY 132 LysGIuPheLeuIyValyThriIyAlaAsnHisIyValaIleMetAsnAlaArgAlaTrpIyS 151
DB 367 AAGCAATTCCTCAAAAAGACCAAGAAAGAACCAAGGCGCATGAAGCGCCGGGCTCGAAG 426
QY 152 ArgTrpCysThriGInIleLeuSerAlaLeuSerPheLeuHisAlaCysSerProProIle 171
DB 427 CGCTGGTCAGCAGAGATCTGTCTGCGCTCAGCTTCTGCACGCTGAGCGCTGAGCCCAATC 486
QY 172 IleHisGIuAAsnLeuThriSerAspThriIlePheIleGInHisAsnGIuIyLeuIleIyS 191
DB 487 ATCCAGGGGAACCTGACAGGACACCATCTTCAATTCAGACACAGGCTTCATCAAGATC 546
QY 192 GlySer-----ValaTrpHisArgIlePheSerAsnAlaLeuArg 204
DB 547 GGCCTCGGTGCTGGCGGGGAGGCGGTGGTGTGGCACCGAATCTTCCAAAT----- 597
QY 205 ProProThraAlaLeuProAspAspLeuArgSerProIleArgAlaGIuArgGIuLeu 224
DB 598 -----GCACCTCCAGATGATCTCGAAGCCCATCCGCGCTGAGCGAGAGGACTT 648
QY 225 ArgAsnLeuHisPhePheProGIuTrpGIuGIuValaAlaAspGIyThriAlaValaAsp 244
DB 649 CGGAACCTGCATCTTCTTCCCGCAGAGTATGAGAGGTGGCGGATGGAGCCGCTGTGAGC 708
QY 245 IlePheSerPheGIyMetCysAlaLeuGIuMetAlaValaLeuGIuIleGInThrAsnGIy 264
DB 709 ATCTTCTCTTGGGATGTGTGCGCTGGAGATGGCTGTACTGAAATCCAGACCAATGGG 768
QY 265 AspThraArgValaIThriGluGlnAlaIleAlaArgAlaArgHisSerLeuSerAspProAsn 284
DB 769 GACACCCGGGTGCACAGAGAGGCAATGCTGCGCGAGGACCTGCTGTAAGTGAACCCAC 828
QY 285 MetArgGIuPheIleLeuCysCysLeuAlaArgAspProAlaArgArgProSerAlaHis 304
DB 829 ATGCGGAGTTCATCTTCTTGTGCTGCTGGCCCGGAACTGCGCGGCTCTTGTCCAC 888
QY 305 SerLeuLeuPheHisArgValaIleuPheGIuValaHisSerLeuIyLeuLeuAlaHis 324
DB 889 AGCTCCTCTTCCACCGCGGTCTCTGAGGTGACCTGCTGAAGTCTCTGGACAGCCAC 948
QY 325 CysPheIleGInHisGInIy-----Leu 332
DB 949 TGTTCATCCAGCACCAATGTAGACAGAGGCGCGCGGTGTGGGAGACCTC 1008
QY 333 MetProGIuAAsnValaGIuGIuIyThriTyraIleMetAspLeuHisAlaValaLeuAla 352
DB 1009 ATGCTTGAGAAATGTGTGTGAGAGAAAGACAAAGCCATGTGACCTGCACGCGGTCTGGCG 1068
QY 353 GIuLeuProArgProArgArgProProLeuGInTrpArgIySerGIuValaSerPheMet 372
DB 1069 GAGCTTCCCGGCGCCGAGGCGCCCGCGCTGACAGTGGCGGTACTCGGAAGTCTCTTCATG 1128
QY 373 GIuLeuAspIySPheLeuGIuAspValaArgAsnGIyIleTyTrProLeuMetAsnPheAla 392
DB 1129 GAGCTGGAACAATTCCTGGAGAGATGTGACGAATGGAATCTACCACTGATGAACCTTTCGA 1188
QY 393 AlaThraArgProLeuGIyLeuProArgValaLeuAlaProProProGIuGIuValaGInIyS 412
DB 1189 GCCACTGACCCCTGGGGCTGCGCCCGTGTGTGTGGCCCAACCCCGGAGAGGTCCAAAG 1248
QY 413 AlaIyThriProThriProGIuProPheAspSerGIuThraArgIyValaIleGInMetGIn 432
DB 1249 GCCAAAGACCCGAGCGCAGAGCCCTTGTACTGTAGACCAAGAAAGTCAATCCAGATGAG 1308
QY 433 CysAsnLeuGIuArgSerGIuAspIyAlaArgTrpHisIleuThriLeuLeuValaLeu 452
DB 1309 TGCAACCTGGAGAGAGAGAGACAAAGCGCGGTGGCATCTCACTCTCTGTGTGCTG 1368
QY 453 GIuAspArgLeuHisArgGIuLeuThriTyraSPLeuLeuProThraSPSerAlaGInAsp 472

RESULT 4
AX056416

Db 1369 GAGAGCCGCTGCACCGGACGCTGACCTGACCTGCTCCCAACGACGAGCCGACGAC 1428

Qy 473 LeuAlaSerGluLeuValHisTyrGlyPheLeuHisGlu-----AspAspArg 488

Db 1429 CTGCGCTCGGAGCTCGGACACTATGCTCTCCACGAGGTGGCGCTGGCGGACGACCGG 1488

Qy 489 MetLeuLeuAlaIlePheLeuGluSerThrPheLeuTyrTyrArgGlyThrGlnAla 507

Db 1489 ATGAGCTGCGCGCTCTCTGAGAGACCTTCTCAAGTACCGTGGAGCCAGGCC 1545

LOCUS AX056416 3304 bp DNA linear PAT 13-JAN-2001

DEFINITION Sequence 60 from Patent WO0073469.

ACCESSION AX056416 GI:12229123

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Plozman, G.D., Martinez, R., Whyte, D. and Sudersanam, S.

AUTHORS Protein kinases

TITLE Patent: WO 0073469-A 60 07-DEC-2000;

JOURNAL Sugen, Inc. (US)

FEATURES

source 1.3304

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 2,3e-196 Length: 3304

Score: 2409.00 Matches: 460

Percent Similarity: 98.72% Conservative: 2

Best Local Similarity: 98.72% Mismatches: 0

Query Match: 90.22% Indels: 6

DB: Gaps: 1

US-10-618-941-99 (1-507) x AX056416 (1-3304)

Qy 40 ArgArgGluGlnValaIleGlnGlnIleAspMetProGlyLeuGlnSerThrPheLeuAlaMet 59

Db 1 CAGAGAGAGAGAGGTAACCAAGGAAATCATGCGAGGCTTCAGAGACCTTCCTAGCCATG 60

Qy 60 AspThrGluGluGlyValaGluValaValTyrAsnGluLeuHisPheGlyAspArgGlyAla 79

Db 61 GACACGAGAGAGGAGGAGTGAAGTGTGTGGAACGAGCTCCACTTCGAGAGCAGAGAGGCC 120

Qy 80 PheAlaIleHisGluGluGlyValaIleGlnThrValaPheGluGlnLeuValaIleAspHis 99

Db 121 TTGCGCGGCGACGAGAGAGATCCAGACGCTGTTCAGAGAGCTGTGTGTGTGTGTGTGTGT 180

Qy 100 ProAsnIleValaIleLeuHisGlyTyrTyrPheLeuAspThrSerGluAlaCysAlaArgVal 119

Db 181 CCGAAGATCGTGAAGTTCAGCAAGTACTGCTGATCTCTGAGGCTGTGCGGAGGCTC 240

Qy 120 IlePheIleThrGluTyrValaSerSerGlySerLeuGlnPheLeuGlyValaIleSer 139

Db 241 ATCTTCATCAGAGT 300

Qy 140 LysAsnHisLeuValaMetAsnAlaArgAlaTyrPheArgTyrCysThrGlnIleLeuSer 159

Db 301 AAGAAACCAAGGCGATGAAGCGCCGCGGCTGTGAAGGCTGTGTGTGTGTGTGTGTGTGT 360

Qy 160 AlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGluValaLeuLeuThrSerAsp 179

Db 361 GCGCTCAGCTTCTCTGACGCTGTGAGCCGCCCAATCATTCACGAGAACTGACCAAGCGAC 420

Qy 180 ThrIlePheIleGlnHisAsnGlyLeuIleLeuValaIleGlySerValTyrHisArgIlePhe 199

Db 421 ACCATCTTCATTCAGACCAACAGGCTCATCAAGATGGCTCCGTGTGGACCGAATCTTC 480

Qy 200 SerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArgAla 219

Db 481 TCCAAAT-----GCACCTTCAGATGATCTCCGAGGCCCATCCGCGCT 522

Qy 220 GluArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyValaIleAsp 229

Db 523 GAGCGAGAGAGAACTTGGAACCTGCACTTCTCCCGCAGAGTATGAGAGGTGGCGCAT 582

Qy 240 GlyThrAlaValaAspIlePheSerPheGlyMetCysAlaLeuGlnMetAlaValaLeuGlu 259

Db 583 GGGACCGCTGTGGAACATCTTCTCTTGGGATGTGTGCGTGAAGATGGCTGTATCTGGAA 642

Qy 260 IleGlnThrAsnGlyAspThrArgValaIleThrGluGluAlaIleAlaArgAlaArgHisSer 279

Db 643 ATCAAGACCAATGGGACACCGCGGTCAACAGAGAGGCCATTCCTCGCGCCAGCACTCG 702

Qy 280 LeuSerAspProAspMetArgGluPheIleLeuGlyCysLeuAlaArgAspProAlaArg 299

Db 703 CTGAGTGACCCCAACATGGGAGTTCATCTTGTGCTGTGCTGCGCGGACCTGCGCGC 762

Qy 300 ArgProSerAlaHisSerLeuLeuPheHisArgValaIlePheGluValaHisSerLeuLys 319

Db 763 CGGCTCTGTGCGCACAGCTCTCTCTTCACCGGCTCTCTTCAGAGTGCATCTGCTGAA 822

Qy 320 LeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetProGluAsnValaIleGlu 339

Db 823 CTCTGTGAGCGCCACAGCTTCATCCAGACAGCACTGCTGAGATGTGTGTGTGTGTGTGT 882

Qy 340 GluLysThrIleValaMetAspLeuHisAlaValaLeuGluLeuProArgProArgArg 359

Db 883 GAGAAACCAAGGCACTGACCTGACGCGGCTCTGCGGAGCTTCCCGCGCCCGCCAGG 942

Qy 360 ProProLeuGlnTyrPheArgTyrSerGluValaIleSerPheMetGluLeuAspLysPheLeuGlu 379

Db 943 CCCCCCTGTGAGGCGGTATCTGTGGAAGTCTCTTATGAGGTGAGCAAAATTCCTGTGAG 1002

Qy 380 AspValaArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGlyLeu 399

Db 1003 GATGTCAAGATGAGATCTTACCACTGATGATGAACTTTGCGAGCCACTGACCCCTGGGGCTG 1062

Qy 400 ProArgValaLeuAlaProProProGluGluValaGlnValaIleAspThrProThrProGlu 419

Db 1063 CCGCGTGTGTGCGCCCAACCGCGAGAGGTCCAAAGGCCCAAGCCCGAGCGCCAGAG 1122

Qy 420 ProPheAspSerGluThrArgGlyValaIleGlnMetGlnCysAsnLeuGluLysSerGlu 439

Db 1123 CCTTTGACTGTGAGACAGAAAGTCAATCCAGATCAGATGCAAGCTTGAGAGAACGAG 1182

Qy 440 AspLysAlaArgTyrHisLeuThrLeuLeuValaIleGluGluAspArgLeuHisArgGln 459

Db 1183 GACAGGCGCGCTGGCATCTCACTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1242

Qy 460 LeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValaHis 479

Db 1243 CTGACCTTACGACCTGTCTCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1302

Qy 480 TyrGlyPheLeuHisGluLysAspArgMetLysLeuAlaAlaPheLeuGluSerThrPhe 499

Db 1303 TATGGCTTCTCTCCACAGAGAGCGAGATGAGCTGCGCTCTGTGAGAGACCACTTC 1362

Qy 500 LeuLysTyrArgGlyThrGlnAla 507

Db 1363 CTCAAGTACCGTGGAGCCAGGCC 1386

RESULT 5

LOCUS HSM801974 3538 bp mRNA linear PRI 18-FEB-2000

DEFINITION Homo sapiens mRNA; cDNA DKFZp434P086 (from clone DKFZp434P086); partial cds.

Db	1183	CGGAGGAGTGTACGGAATGGAATCTTACCACTGAGTAACCTTGGACCACTGCACCTCTG	1242		
QY	338	GLYLEUPProArgValIleuAlaProProProGluIuValGlnLysAlaLysThrProThr	417		
Db	1243	GGGCTGCCCGCTGTCTGGCCCCACCCCGGAGGAGGTCCAAAGGCCAAGCCCCGAGCG	1302		
QY	418	ProGluProPheAspSerGluThrTrpLysValIleGlnMetGlnCysAsnLeuGluArg	437		
Db	1303	CCAGAGCCCTTTGACTCTGAGACCGAAGGTCAATCGAGATGCACTGCAACCTGGAGGA	1362		
QY	438	SerGluAspLysAlaArgTrpHisLeuThrLeuLeuValIleGluAspArgLeuHis	457		
Db	1353	AGCGAGACAGAGGCCCGCTGGCATCTCACTGCTTCTGGTGTGGAAGACCGGCTGCAC	1422		
QY	458	ArgGlnLeuThrTrpAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeu	477		
Db	1423	CGCGAGCTGACCTTACGACTGCTGCCAACGACAGCGCCAGAGACTTGCTCGGAGCTC	1482		
QY	478	ValHisrTrGlyPheLeuHisGluAspAspArgMetLysLeuAlaPheLeuGlnSer	497		
Db	1483	GTGCCTATGTGCTTCTTCACACGAGGACGACCGGATGAAGCTGGCGCTTCTGGAGAGC	1542		
QY	498	ThrPheLeuLysTrpArgGlyThrGlnIle	507		
Db	1543	ACCTTCTCAAGTACCGTGGGACCCAGGCCC	1572		
RESULT 6	AR310264	981 bp	DNA	linear	PAT 12-JUN-2003
LOCUS	AR310264	Sequence 9 from patent US 6558903.			
ACCESSION	AR310264				
VERSION	AR310264.1	GI:31703044			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 981)				
AUTHORS	Hodge, M.R.				
TITLE	Kinases and uses thereof				
JOURNAL	Patent: US 6558903-A 9 06-MAY-2003;				
FEATURES	Millennium Pharmaceuticals, Inc.; Cambridge, MA				
Source	Location/Qualifiers				
	1..981				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ALIGNMENT SCORES:					
Alignment Scores:	4,12e-129	Length:	981		
Pred. No.:	1617.00	Matches:	312		
Score:	96.89%	Conservative:	0		
Percent Similarity:	96.89%	Mismatches:	3		
Best Local Similarity:	60.56%	Indels:	7		
Query Match:	6	Gaps:	1		
DB:					
US-10-618-941-99 (1-507) x AR310264 (1-981)					
QY	1	MetaIaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluArgGluAsp	20		
Db	35	ATGGGGGGCCCGGAGCCCGGCCGAGGGCGGCCGGGAAACGGAGCGGAGCCGGAGAGAC	94		
QY	21	GluSerGluAspGluSerAspIleuGluGluSerProCysGlyArgTrpGlnLysArg	40		
Db	95	GAGAGCGAGAGCGAGAGGACATCTCGTAGGAAAGCCCGTGTGCTGCTGGCAAAAGCA	154		
QY	41	ArgGluGlnValaGlnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp	60		
Db	155	CGGAGCGAGTAAACCAAGGAAACATGCGAGGCGTTGAGACACCTTCTAGCCATGAC	214		
QY	61	ThrGluGluGluValaGluValaTrpArgGluLeuHisPheGlyAspArgLysAlaPhe	80		
Db	215	ACGGAGGAGGGGGTGAAGTGTGTGGAAACAGCTCCACTTGGAGACAGGAGGCTTTC	274		

OY		81	AlaIahAsgluGluLysIleGIlnThrValPhegluglInLeuValIleuValAspiApro	100
Db		275	GCGGGCAGAGAGAAATCATCAACCGGTTCGAGCGACTGGTCTTGTTGGACCAACCG	334
OY		101	AenIIeValIyLEuNIeLyTyTTrpLeuApThSerSgiuaIaCySaIaArgValIle	120
Db		335	AACATCGTAGAGTTGCACAAGAATCGGTGGATACCTTGAGGCCGTGGCGGAGTATC	394
OY		121	PheIIeThrgIuTYrValISerSergIyerIEuleylseginPheLeuVlygThrlYelys	140
Db		395	TTCATCACAGAGTACGTGTTCATCAGGCGGCTCAAGCAATTCCCAAAMAAACCAAGAG	454
OY		141	ASnHIslySaImetASnaIaARqAlatRpyVaArGTAPCyStrngInIleuSERaIa	160
Db		455	AAACCAGAGGCCATGAACCCCCGGGCTGGAAAGCGCTGTGCAGCAATCCTGICTCG	514
OY		161	IeuSeRPHeuHIaIaCYseSeRPProIleIlleNIsgIyaSnIleuThrSerApThr	180
Db		515	CTCAGCTTCCTGCAAGCCTGCAGGCCCCCAATCATCCAGGGAACITGACAGCGAAC	574
OY		181	IlePheIIegInHIsaNglyLeuNILEySIleglYserValITPhISarGIlePheSer	200
Db		575	ATCTTCATTCAGCACAAGCGGCTCATCAAGATCGGCTCGTGGCAGCCGAATCTTCTCC	634
OY		201	ASnaIaleuARGProPtoThrAlaIeunProsaPheIueaRGserProIleaRqAlaglu	220
Db		635	AAAT-----GCACTTCAGATGATCTCCGAAGGCCCATCGGCGCTAG	676
OY		221	ArgIugIuleuAARGsnIleuHIsPhePheRPProgluTYrGIyguIalaaSPsily	240
Db		677	COAGAGGAATCTTCGGAACCTGCACATCTTCCCCCAAGATGAGAGAGTGGCGGATGG	736
OY		241	ThraIaValaIpIlePheSerPheglYmeCYsaIaleuGIumeCIaIaValleuGlulle	260
Db		737	ACCGGTGGAACATCTTCTTGTGGATGTGTGCGCGGAGATGGCTGTACTGAATATC	796
OY		261	GIlnThraNGIYAaPThrArgValIThrgIugIuaIaiIeaIaRqAlaargHIsSerIeu	280
Db		797	CNAGCAATGGGGACACCCGGGTCTACAGAGAGGCCATTTGTCGGCCAGGACATCGCTG	856
OY		281	SerApProaAmEtargIuPheIleuCYsaIeunaIaARGspProIlaARG	300
Db		857	AGTGAACCCACATGGCGGAGTTCATCTTGTGCTGGCGCGGAGACCTGCCGNCGG	916
OY		301	ProSeRaIahIsSerIEuleuPheHIs-ArgValleuPhegluValHIsSerIEuleyle	320
Db		917	CCCCTGTCCACAGCCTCTCTTCCACACGCGTGTCTTNGAGGTGACATCGCTGAAGCT	976
OY		320	uIeu 321	
Db		977	CCTG 980	
RESULT 7				
LOCUS	ARE40307		981 bp	DNA linear PAT 20-APR-2005
DEFINITION	Sequence 9 from patent US 6858418.			
ACCESSION	ARE40307			
VERSION	ARE40307.1		GI:62774642	
KEYWORDS	.			
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 981)			
AUTHORS	Hodge,M.R.			
TITLE	Kinases and uses thereof			
JOURNAL	Patent: US 6858418-A 9 22-FEB-2005;			
FEATURES	Milennium Pharmaceutical, Inc.; Cambridge, MA			
SOURCE	location/Qualifiers			
	1..981			
	/organism="unknown"			
	/mol_type="genomic DNA"			

AUTHORS

Birtten, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Bogunlavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dattellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hages, B., Hall, J., Horton, L., Holme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karacas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramamany, U., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, C., Spence, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talama, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL

COMMENT

Submitted (02-APR-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 2, 2003 this sequence version replaced gi:28827929.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information

Center project name: L23072

Center clone name: 299_M_14

FEATURES
source

Only the middle 92.3 Kilobases of this clone are being submitted.
The remainder overlaps either accession number AC109322 [WICGR project L23183]
or accession number AC105219 [WICGR project L23081].

Location/Qualifiers

1. 32291

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="8"

/map="8"

/clone="RP11-299M14"

/clone_lib="RP11-299M14 Human Male BAC"

repeat_region

1349..1371

/rpt_family="GC_rich"

repeat_region

2401..2449

/rpt_family="T-rich"

repeat_region

complement(3976..4326)

/rpt_family="L1MD2"

repeat_region

complement(4473..4777)

/rpt_family="AluSx"

repeat_region

4787..4920

/rpt_family="(TACA)n"

repeat_region

complement(4924..5201)

/rpt_family="AluY"

repeat_region

complement(5208..5350)

/rpt_family="L1MD2"

repeat_region

5369..5596

/rpt_family="AluSx"

repeat_region

complement(5719..6043)

/rpt_family="L1MD2"

repeat_region

complement(6044..6347)

/rpt_family="AluSx"

repeat_region

complement(6348..6820)

/rpt_family="L1MD2"

repeat_region

complement(6953..7076)

repeat_region	/rpt_family="L1MD2"	7326..7349	/rpt_family="(CAAA)n"	complement(7377..7458)	repeat_region
repeat_region	/rpt_family="L1MD2"	8091..8207	/rpt_family="MIR"	8208..8505	repeat_region
repeat_region	/rpt_family="AluSg"	8506..8580	/rpt_family="MIR"	complement(8785..9127)	repeat_region
repeat_region	/rpt_family="AluSg"	9225..9304	/rpt_family="ATGCTG)n"	9420..9660	repeat_region
repeat_region	/rpt_family="AluSg"	9661..10614	/rpt_family="L1MD2"	10615..10690	repeat_region
repeat_region	/rpt_family="AluSg"	10697..10994	/rpt_family="AluY"	11474..12045	repeat_region
repeat_region	/rpt_family="AluSg"	complement(12065..12372)	/rpt_family="AluY"	12379..12519	repeat_region
repeat_region	/rpt_family="AluSg"	complement(12918..13211)	/rpt_family="AluSx"	complement(13213..13346)	repeat_region
repeat_region	/rpt_family="AluSg/x"	complement(13363..13736)	/rpt_family="MER87"	complement(14215..14374)	repeat_region
repeat_region	/rpt_family="MLT1P2"	14698..15000	/rpt_family="AluO"	complement(15346..15495)	repeat_region
repeat_region	/rpt_family="MLT1C"	complement(15500..16464)	/rpt_family="L1MD2"	complement(16472..16738)	repeat_region
repeat_region	/rpt_family="L1MD2"	16783..16874	/rpt_family="L1MD2"	complement(16869..17347)	repeat_region
repeat_region	/rpt_family="L1MD2"	complement(17349..17752)	/rpt_family="MER77"	17753..17956	repeat_region
repeat_region	/rpt_family="MER30"	complement(17957..18089)	/rpt_family="MER77"	complement(18519..18684)	repeat_region
repeat_region	/rpt_family="G-rich"	19020..19200	/rpt_family="G-rich"	19204..19235	repeat_region
repeat_region	/rpt_family="GC_rich"	19295..19344			repeat_region

Alignment Scores:

Pred. No.:	3.9e-126	Length:	92291
Score:	1612.00	Matches:	498
Percent Similarity:	28.84%	Conservative:	1
Best Local Similarity:	28.79%	Mismatches:	8
Query Match:	60.37%	Indels:	1231
DB:	8	Gaps:	17

US-10-618-941-99 (1-507) x AC105049 (1-922291)

[illegible]

Db	20568	ACAAAGCCATGAAGACGCCCGGAGTATGGGGAGCGGGGCTGGGGACAGCCACGGGGACAGAGACGG	20627
QY	148	-----	148
Db	20628	GGTTGGGGCAGCCTCGGGGACTGGGATGGGTAGAGGGGGGTGCCCGCGGCTCGGACAGGGG	20687
QY	148	-----	148
Db	20688	CTGGGCGAGGATCGGGGGCGGGCTCGAGAGCCAGCGGCTCTCTGCGCCACCCGAC	20747
QY	149	-----	149
Db	20748	CGACGAGTGTGCTGCCGCCGACGAGCCTTGAAAGCGCTGGTGCACGAGATCTGTCTG	20807
QY	160	1a-----	160
Db	20808	CGCTCAGGTAGGGGCGTGGGCTGGGCGCTGGGGCGGAGCCAGAGCCCCCATCTCTCAG	20867
QY	161	-----	161
Db	20868	AGCCTCTGACCGGCTCTCTCTACTCTCCCTCCCAAGTGTGAGCTTCTTGACGCC	20927
QY	167	CysSerProRoblleiIehIsigIyaenLeuThrSerAspThrIlePhelIegInIshsn	186
Db	20928	TGCAGCCCCCATTCATCCAGGAACTGACACGAGACACCATCTTCATTCAGACAAAC	20987
QY	187	GIyLeuIleIyIsIegIySer	193
Db	20988	GGCCTCATCAAGATCGGCTCGGCTGGGCGGGGCGAGGCTGGGTGGGGAAACGGGGCAG	21047
QY	193	-----	193
Db	21048	GGGAAGGTGCTGAGAGCCAGGACAGAGAGTTCTGATGCCACAGCTAACTCCATC	21107
QY	194	-----	194
Db	21108	TTTCTCCCGGCCCTTGGGAGTGTGGACCGGAATCTTCTCAATGG-TAAGTCGGACTTT	21166
QY	203	-----	203
Db	21167	GGGGGGAACGGGGTCTGAGAGGTCCCGCCGCTCAGCTGCGCCCTCCACAGCACT	21226
QY	209	uProAspAspLeuArgSerProIleArgIaIuArgIuLeuArgAsnLeuHisph	229
Db	21227	TTCAGATGATCTCGAAGCCCCCATCCGCGCTGAGGAGAGAACTTCGAACTTGACTT	21286
QY	229	ePheProGluIuTyrlY	235
Db	21287	CTTCCCCCAGAGATGAGAGGTAGTGTGTGGTCCGGGTGCCCGGCTTACCTGTGC	21346
QY	236	-----	236
Db	21347	GCTTCTCTCTAGGTGTCTGTGTCTCTTCTGTTCAGAGGTGGCCGAGTGGACGGTGT	21406
QY	243	IAspIlePheSerPheGlyWetCysAlaIeuGlu	254
Db	21407	GGAATCTTCTCTTGGAGTGTGTGTGCGCTGGA-GGTACTGCCCGCCCTCCACTTGT	21465
QY	254	-----	254
Db	21466	GCTGTGCGCTCTCTAACTGTGCCCTGAGCTGACTGATGTGTGCTCTTCCCATGC	21525
QY	255	-----	255
Db	21526	CTGACCTGGCTCTGTGCCCGAGATGGCTGTACTGAGAAATCCAGACAAATGGGAGACCCG	21585
QY	267	GValThrGluGluIaIleIaArgIaArgHisSerLeuSerAspProAsnMetArg--	286
Db	21586	GGTACAGAGGAGGCACTTGTCTGCCCGCAGGACCTGTGTAAGTGAACCCCAATGTGGGT	21645
QY	286	-----	286

Accession	Gene	Chromosome	Position (kb)	Orientation	Sequence	Length (bp)
AY001	428	1	174949	+	AC116487	428
AY002	23862	1	174949	+	AC116487	23862
AY003	428	1	174949	+	AC116487	428
AY004	23922	1	174949	+	AC116487	23922
AY005	428	1	174949	+	AC116487	428
AY006	23982	1	174949	+	AC116487	23982
AY007	428	1	174949	+	AC116487	428
AY008	24042	1	174949	+	AC116487	24042
AY009	429	1	174949	+	AC116487	429
AY010	24102	1	174949	+	AC116487	24102
AY011	433	1	174949	+	AC116487	433
AY012	24162	1	174949	+	AC116487	24162
AY013	445	1	174949	+	AC116487	445
AY014	24222	1	174949	+	AC116487	24222
AY015	446	1	174949	+	AC116487	446
AY016	24282	1	174949	+	AC116487	24282
AY017	461	1	174949	+	AC116487	461
AY018	24342	1	174949	+	AC116487	24342
AY019	467	1	174949	+	AC116487	467
AY020	24402	1	174949	+	AC116487	24402
AY021	471	1	174949	+	AC116487	471
AY022	24462	1	174949	+	AC116487	24462
AY023	485	1	174949	+	AC116487	485
AY024	24521	1	174949	+	AC116487	24521
AY025	486	1	174949	+	AC116487	486
AY026	24581	1	174949	+	AC116487	24581
AY027	501	1	174949	+	AC116487	501
AY028	24641	1	174949	+	AC116487	24641

REFERENCE
AUTHORS
2 (bases 1 to 174949)
Birtren,B., Litron,L., Nusbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Glnde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Labouque,R., Lamazares,R., Landers,T., Lehoczeky,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Maynor,J., Margutsu,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Minova,T., Mlenka,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Pollard,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rossetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tophan,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (28-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 174949)
Birtren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Archchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearlano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R., McClean,C., MacDonald,P., Mayor,J., Manning,T., Matthews,C., McCarty,M., Meldrum,J., Meneus,L., Minova,T., Mlenka,V., Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunhthang,P., Pierre,N., Rachupka,A., Ramsamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J., Tophan,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (02-NOV-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 2, 2005 this sequence version replaced gi:71067273.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
Genome Center
Center: Broad Institute of MIT and Harvard
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@broad.mit.edu
Project Information
Center project name: L23947
Center clone name: 428_1_9

NOTE: This is a 'working draft' sequence. It consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

QY 244 AspIephSerPheGlyMetCysAlaLeuGlu----- 254
 Db 7337 GACATCTTCTCTTCGGGATGTGTGCACTGGA--GTACTGTCCACTCTCTCATAGCCC 7395
 QY 255 -----MetAlaVal 257
 Db 7396 TGTCCAGGGTCTACGGCATCCACCAATTCCAATTGTCTCCCTGACCCAGATGGCGTGA 7455
 QY 258 LeuGluIleGlnThrArgValAspThrArgValThrGluGluAlaIleAlaArgAlaArg 277
 Db 7456 CTGAGATTCGACCAACCGGGGATACCAAGATCAAGAGAGGCGCATGCTCGAGCGCAGG 7515
 QY 278 HisSerLeuSerAspProAspMetArg----- 286
 Db 7516 CACTCACTAGTGAACCCCAACATGCGGGTAGAGCACTTACCTTGCCCTTGGGGCTGGC 7575
 QY 286 ----- 286
 Db 7576 TGCAGGTGCTTCAAGCTTACCTTGAAGTGTCTTAGGAGGCGCAAGGGGCGAGCTGGTG 7635
 QY 286 ----- 286
 Db 7636 GTAATCTGAGAGGGGCGAGCTGTGTCAGTACATCTGAGAGTGAAGAGAGACTAAGGC 7695
 QY 286 ----- 286
 Db 7696 CCACCTGAAAGAAAGACTCTTCCAGAAAGCTCACAGACTCCACAAAGGTGAATGAAAGTAG 7755
 QY 286 ----- 286
 Db 7756 GCATACGTGTTTGAAGCCCAAGGCTTGCTGTCTCAAGTGTCCAGAAATCAGATCTTGAG 7815
 QY 286 ----- 286
 Db 7816 CTGCTAATGGAGTGAATCTCTGTCTCATATCAAGAGAGCTTTCAGACATCCTGAG 7875
 QY 286 ----- 286
 Db 7876 GTGGAGATGTGTGATTTGGACCGGTGGGGTCTAAGTCTTGAAGACCTCTCTAGAG 7935
 QY 287 -----GluPheIleLeuCy 291
 Db 7936 CGAGACTTTTCAGACTATCACTGTCTTCTTGTTTCAACCTAAGAAATTCATCTCTTC 7995
 QY 291 sCySleuAlaArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgVal 311
 Db 7996 CTGCTGTGGCCCGGACCTGTGCGCGACCTTCAGGCCACAACTCTCTTCCACCGAGT 8055
 QY 311 IleuPheGluValHisSerLeuIysLeuLeuAlaAlaHisCysPheIleGlnHisGln-T 331
 Db 8056 GCTCTTTGAGGTGCATCTGCTGAAGCTGTGGAGCTCACTGCTTCAATCCAGCACAGTG 8115
 QY 331 Yr----- 331
 Db 8116 TGAGGGGCAAGAGGCGGGGTGGGTCAAGATCTGGGGTCAAGTGTGAGAGGTCCTTT 8175
 QY 332 -----LeuMetP 334
 Db 8176 GCAGAGCCCTTAGATTTGGGTGCATCCCATAGACACTTGTGTCTTCCACAGACTCATGG 8235
 QY 334 roGluIleuValIleGluGluIleThrIysAlaMetAspLeuHisAlaValLeuAlaGlu 354
 Db 8236 CTGAGATGTGTGAGAGAAAGACCAAGGCCATGACCTTCATGCAAGTTTGGCTGAG 8295
 QY 354 euProArgProArgArgProProLeuGlnIleArg----- 365
 Db 8296 TGCCTGAGGCCCATGAGACCCCAATGACGTGGCG--GTAGATAGTGGCAACGAGAGCTTC 8354
 QY 365 ----- 365
 Db 8355 CCTTCTAGCTGTGGCTTAGGCTTGGCCCTGACATCTCATATGTAACCTTCCACACCAT 8414
 QY 366 -----TyrSerGluValSerPheMetCyluLeuAspIysPheLeuGluAspVal----- 381

Db 8415 CTAGGACTCAGAGGTCTCTCTTGTGGAGCTGGACAAATTCTTAGAGATGT--CAGGTGA 8473
 QY 381 ----- 381
 Db 8474 GAGCTGCATGAGACTGAAGAAAAGCACTATCAGGCTGAGCCCCACCTAGACTACCA 8533
 QY 382 -----ArgAsnGlyIleTyrProLeuMetAspPheAlaAlaThrAspP 396
 Db 8534 CCCCCTGTCTATCCCAAGAAAGGATCTATCTCACTGATGAACCTTGGCGCTCGGC 8593
 QY 396 roLeuGlyLeuProArgValIleuAlaProProGluGluValGlnIysAlaIleThrP 416
 Db 8594 CTTGGGGCTTCCCGTGTGTGTGGCCCAACCCCAAGAGAACCCCAAGGCCAAAACTC 8653
 QY 416 roThrProGluProPheAspSerGluThrArgValIle----- 429
 Db 8654 CAACGCCAAGAACCTTTGACTCGSAGACCAAGAAAGGTGAG--TCCCACTCTGTGTACT 8712
 QY 429 ----- 429
 Db 8713 GTGCTGTCTTAGAGTTGACCAATCAATGAGAAATGGAATGAGATCTTCTCTGGGGA 8772
 QY 429 ----- 429
 Db 8773 CCGGCTCTCCCATCCCGGTTCTCCCCAGGCTGCGAAGACCAATCCAGCTGACCC 8832
 QY 429 ----- 429
 Db 8833 CTTTGTGAAGCTGGAACCGGACCGACACATATTTCCCAAGAGCTTCTGTCTTAC 8892
 QY 429 ----- 429
 Db 8893 AGTCTTCCCAATCCGTTGATGTCTCTTCCAGAGGCTTGGCTAGCTGTGTCCTCC 8952
 QY 429 ----- 429
 Db 8953 GTGTGATCCAGAGCATTTGCTCCGGTCTCCCTTCTCATAGCTTCTGTCTCTCAATA 9012
 QY 429 ----- 429
 Db 9013 CACCTTAATTCAAATTCAGATGCTACATGAGCTGCCAGAGAACTAGTGAATCAG 9072
 QY 429 ----- 429
 Db 9073 CCTAAGCCCAAGCTCTGTCTCTCTCTGTGGCACTTGTGTTTCTGTACAGTCAC 9132
 QY 429 ----- 429
 Db 9133 CCGGGGGGCAAGGTTTCTTGTCTTCCAGGTTCTCTTCTGTCTGTCTCTGC 9192
 QY 429 ----- 429
 Db 9193 CTCCCATCCCTCTCTGGAATGTATTCATATTCAGAGCAACCAACGAAGCTGTCT 9252
 QY 429 ----- 429
 Db 9253 TGAAGCTTCCCTAATCTCAGGGCTGTGTGAGTAAGAGTGTCTTTCGACTTCCCTTG 9312
 QY 429 ----- 429
 Db 9313 GAGACATTTACATCTTGGGTCTGTGTCTCTCTTCTGGGCTCTCTGTAAAGCCGTG 9372
 QY 430 -----GlnMetGlnCysAsnLeuGlu 437
 Db 9373 CTCTGAGCTAATACACTTCCCTGCGCAAGGTGTTCAGATGCACTCAACTGGA 9432
 QY 437 rGSerGluAspIysAlaArgTyrHis----- 445
 Db 9433 GAAGGAGGACAAAGCTCGGTGCACTGAGCAAGGGGCTGTGAGACTAGGGAAGAT 9492
 QY 446 -----LeuThrLeu 449

REFERENCE	TITLE	JOURNAL	AUTHORS
9493	GGGAGCAATCAACTCTGGGCGCGCTGACACAGCCCTTCGCTGAGCTTACTCTGC		9552
9499	euLeuValleuGlunAphAryleuH1eAryglInLeuTh1TyAryleuLeuProthr		467
9553	TCCTGGTCTTGAAGACCGGCTACATCGGACAGCTACCTATGATCTGCTCCAG-TAG		9611
9467	-----		467
9612	CTGGTAGTGGGTCAGGTGTGGGTGAGGCTCAGGCTCAGGACCTAGGCTATGTCCTACTATT		9671
9468	-----AppSer1aGlnAphLeuA1aSerGlnLeuValH1st		480
9672	TGCTGTTGTTCTCTTCAAGGACAGAGCCAGAGCTGCTGCTGACTGACTGACTT		9721
9480	YTGlyPheLeuH1aGlu-----		485
9732	ATGGCTTCTGACAGA-GATGGCGGCGTGGCTTGGTGTGGAAAGTGACGGGTCTGC		9790
9486	-----AspAphAryleuH1aGlu		491
9791	CGAGTGGCGGCTACTACCACTGCTTCCCTCATCTCAGTCACTGACGAGATACAGGACAAAGC		9850
9491	eulA1a1AphLeuGluSerThyPheLeuTy1TyAryGlyThGlnA1a		507
9851	TAGCAGCTTCTTCTGAGACACACTTCTTCAAGTACCGAGGACGACGACG 9900		
RESULT 10	AC110211/C	202844 bp	DNA linear ROD 05-NOV-2004
LOCUS	Mus musculus chromosome 15, clone RP23-192821, complete sequence.		
DEFINITION	AC110211		
ACCESSION	AC110211.16	GI:55416090	
VERSION	HTG.		
KEYWORDS	Mus musculus (house mouse)		
SOURCE	Mus musculus		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 202844)		
AUTHORS	Bitren,B., Nusbaum,C. and Lander,E.		
TITLE	Mus musculus chromosome 15, clone RP23-192821		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 202844)		
AUTHORS	Bitren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baetsen,V., Boguski,K., Bouckgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Charzo,R., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gaidyna,S., Glade,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kelle,C., Lacroque,K., Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C., McDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menusz,L., Minova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausn,K., Subramanian,A., Talamas,J., Teeffaye,S., Theodore,J., Topham,K., Traversman,M., Travis,N., Triggilio,J., Vassiliou,H., Viell,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (10-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 202844)		
AUTHORS	Bitren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N., Anderson,M., Anderson,S., Archchi,H.M., Barna,N., Baetsen,V., Bloom,T., Boguski,K., Bouckgalter,B., Camarata,J., Chang,J., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Corum,B.,		

TITLE	JOURNAL	AUTHORS
Submitted (05-OCT-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 202844)		Darvillano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., McLaren, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, U., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnipack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, D., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.
TITLE	JOURNAL	AUTHORS
Submitted (05-NOV-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA 4 (bases 1 to 202844)		Biren, B., Nusbaum, C., Lander, E., Abouelell, A., Allen, N., Anderson, M., Anderson, S., Archchi, H.M., Barua, N., Bastien, V., Bloom, T., Boguskiy, L., Bokhalter, B., Camarata, J., Chang, J., Chepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hago, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Liu, X., Lui, A., Mabbitt, R., MacLean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrim, U., Menus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nguyen, T., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schnipack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, D., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.
COMMENT	JOURNAL	
All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html		
FEATURES	SOURCE	
Location/Qualifiers		
1..202844		
/organism="Mus musculus"		
/mol_type="genomic DNA"		
/db_xref="taxon:10090"		
/chromosome="15"		
/map="15"		
/clone="RP23-192B21"		
/clone_lib="RPCI-23 Female Mouse BAC"		
/complement(1609...11757)		
/rpt_family="Bl_MM"		
7176...7216		
/rpt_family="(CCG)n"		
9059...9095		
/rpt_family="(GA)n"		

```

repeat_region      11282..11303
/rpc_family="AT_rich"
repeat_region      11374..11512
/rpc_family="HAL1"
repeat_region      11589..11721
/rpc_family="B1_MM"
repeat_region      11799..11947
/rpc_family="B1F"
repeat_region      12054..12092
/rpc_family="(CA)n"
repeat_region      12349..12472
/rpc_family="B1_MM"
repeat_region      complement(12490..12541)
/rpc_family="B1_MM"
repeat_region      12558..12598
/rpc_family="Lx4"
repeat_region      12652..12978
/rpc_family="Lx4"
repeat_region      complement(13072..13151)
/note="single clone coverage"
repeat_region      complement(13142..13146)
/note="30 Qual SINGL region"
repeat_region      13658..13820
/rpc_family="RSINE1"
repeat_region      13825..13867
/rpc_family="(CA)n"
repeat_region      13876..13952
/rpc_family="HAL1"
repeat_region      14121..14197
/rpc_family="B1_MM"
repeat_region      14199..14665
/rpc_family="MT2B"
repeat_region      14826..15002
/rpc_family="MT2B"
repeat_region      14892..15021
/rpc_family="B1_MM"
repeat_region      15022..15043
/rpc_family="AT_rich"
repeat_region      15241..15692
/rpc_family="MT1B3"
repeat_region      15693..15723
/rpc_family="(TG)n"
repeat_region      15724..15776
/rpc_family="MT1B3"
repeat_region      complement(15890..16025)
/rpc_family="L2"
repeat_region      20971..21109
/rpc_family="C-rich"
repeat_region      22357..22540
/rpc_family="B3A"
repeat_region      22726..23017
/rpc_family="L2"
repeat_region      23800..23946
/rpc_family="MT1F2"
repeat_region      24599..24743
/rpc_family="B1_MM"
repeat_region      complement(24761..24909)
/rpc_family="MT1A"
repeat_region      complement(24977..25152)
/rpc_family="B3"
repeat_region      25153..25466
/rpc_family="ORR1A3"
repeat_region      25509..25603
/rpc_family="PB1D10"
repeat_region      complement(25680..25823)
/rpc_family="B1_MM"
repeat_region      25853..26024
/rpc_family="B2_Mm1"
repeat_region      27294..27381
/rpc_family="(GA)n"
repeat_region      complement(27387..27494)
/rpc_family="PB1D10"
repeat_region      27735..27802

```

```

repeat_region      /rpc_family="(GAAA)n"
repeat_region      complement(28677..28873)
/rpc_family="URR1A"
repeat_region      29722..29903
/rpc_family="U1M2"
repeat_region      30215..30415
/rpc_family="B3"
repeat_region      30422..30447
/rpc_family="AT_rich"

Alignment Scores:
Pred. No.:      1,56e-121      Length:      202844
Score:          1563.50      Matches:      469
Percent Similarity: 32.50%      Conservative: 12
Best Local Similarity: 31.69%      Mismatches: 24
Query Match:      58.56%      Indels:      984
DB:               9      Gaps:      19

US-10-618-941-99 (1-507) x AC110211 (1-202844)

QY      1      MetalAlaProGluProAlaProArgArgAlaArgLuarGluArgLuarGluAsp 20
         |||
Db      21043 ATGGCGGCCCGGAGCGCGCGCGCGAGAGAGGC-----CGGAGCGCGAGCGGAGAGAC 20990
         |||

QY      21      GluSerGluAspGluSerAspAlaLeuGluGluSerProCysGlyArgTArgLysArg 40
         |||
Db      20989 GAGAGCGAGAGCGAGAGCGAGATCTGAGAGAGCCGAGTGGCCGCTGGAGAGCGG 20930
         |||

QY      41      ArgGluGluVal-----
         |||
Db      20929 CGGAGAGCAGGT-GGGCAGCGGTGCGGGTGAACCGCGGTCCCTAGAGCTAGAG 20871
         |||

QY      44      -----
Db      20870 CCATGGGATCCGTGGGGCGGAGCCAGTATAGAGGTCTCCCGAGCGCCTGGCCTTGA 20811
         |||

QY      44      -----
Db      20810 GGCCCTGGATCCTTTGGCCCGGAGCCCTGTGTAGAAACCCCATTCCTTGAACACCGG 20751
         |||

QY      44      -----
Db      20750 TGTGATAGTGGGACCACTCCAGGTGCTGTGATCCCTACTACCGGTGTGGGATCTC 20691
         |||

QY      44      -----
Db      20690 CCCAAGCTTAAGAACTGCGTACCTTGTGGCCAGCCCTGGAGACCACTTCTCCCC 20631
         |||

QY      44      -----
Db      20630 TGGGTTCTGATTCACAGCTTACCTCTCAGCCCTCCACTGAGCTGGGAGAGCTAGC 20571
         |||

QY      44      -----
Db      20570 AGCAGATTACAGGCTGGGGGAGAGGTGTGGGGAGACCAAGTCCACACCTGGGTCTATT 20511
         |||

QY      45      -----
Db      20510 CGTATCAGACACACATTCCTTCCACCTAGGTGAACAGGAAACATGCCCGGAGATCCAG 20451
         |||

QY      54      SerThrPheLeuAlaMetCAspThrGluGluValGluValValLTPAengLysHis 73
         |||
Db      20450 AGCAGCTTCTGGGCAAGAGACCGAGAGAGGTGTAGAGTGTGTGAGAGAACTTAC 20391
         |||

QY      74      PheGlyAspArgLysValaPheAlaAlaHisValu-----
         |||
Db      20390 TTGGGGAGTGAAGAGGCTTTGGCGGCCATGA-GTGAAGCCTGTATACCCGCTTGC 20332
         |||

QY      85      -----
         |||
Db      20331 CCCATACCGTTGATTTGGCCCAACCAATCCTTCTCCAGGAGAGATCCAGACCATG 20272
         |||

QY      91      PheGluGluLeuValLeuValAspHisProAsnLeuValLysValLysValTyrTyrPhe 110
         |||

```


Db 20271 TTTGACAGAGCTGCGCTGGTGAGCAATCCCAACATCTGCAAGCTGCACAGATATGGCTG 20212
QY 111 AspThrSerGluIle----- 115
Db 20211 GAGGCTCTTAGGC-CCGCCGAGGGGTAGCTAGGGCGGGTAAAGGCGGAGTGAAGCC 20153
QY 116 -----CysAlaArg 118
Db 20152 ACAGGCTGGGTGGGAGCATGGAGACCACCGCAACGGCGGAGCCCGCTCCCGCAG 20093
QY 119 ValIlePheIleThrGluTyrValSerSerGlySerLeuIleGlnPheLeuIleValThr 138
Db 20092 GTCATCTTCATCACAGAAATACGTGTGCTGCGCAGGCTCAAGAGATTCTCAAAAGAGCC 20033
QY 139 LysLysAsnHisIleValIleMetLeuAlaArg----- 148
Db 20032 AAGAAAGAACCAAGGCGCATGAACGCCCGGGTATGGGAGTGGGCTGGGGCATGGAGGA 19973
QY 148 ----- 148
Db 19972 GTCCACAGAACTTAGGAGCAGTATGGGGCGGGGAGCTCCGAAAACCTCGGTTCAAGA 19913
QY 148 ----- 148
Db 19912 GTGGGCTAGCGGGGCTCCGAGGCTCAGACGCTTTTGTGCCCAACCGACGAGATC 19853
QY 149 -----AlaTrpLysArgTrpCysThrGlnIleLeuSerAlaLeu----- 161
Db 19852 GTGCGTCCGCTGCAGGCGCTGGAAAGCGCTGGTATAGCAGATCTGTCTGCACCTCAGGTG 19793
QY 161 ----- 161
Db 19792 AAGAGCTGGACCTTGCCACGCGCTGGCCCCCTGCTGCACCTTTCCGCTCTCTGAC 19733
QY 162 -----SerPheLeuHisAlaCysSerProPr 170
Db 19732 CCGGTTCTTCCTGCTCTCATCTCTATGCTTGCAGATTTTTCGACGCGCTGCAGTCCCC 19673
QY 170 ollelleHleHleGlyAsnLeuThrSerAspThrIlePheIleGlnHisAsnGlyLeuLeuIle 190
Db 19672 CATCATCCACGGGAACTTGAACGACACCATCTTCATTCAGCACAATGGCCCTCATCA 19613
QY 190 silleGlySer----- 193
Db 19612 GATTCGGCTCCGGTGTGGCGGAGCGGGGACAGGGGCGAGGGGCTGGGGAACCTGC 19553
QY 193 ----- 193
Db 19552 GGGAGCCCTGACAGAGACTGGCCAGTCTCCCGCCCACTGACATCTGTCTCTCTC 19493
QY 194 -----ValTrpHisArgIlePheSerAsn----- 201
Db 19492 TGCCTGTGGCTGTTCAGTGTGTGTAACCGCATCTTCTCAATGGTGTGTGGGCTCTCGGG 19433
QY 202 -----AlaLeuArgProPro-----ThrAlaLeuPro 211
Db 19432 GTGTGGGGACATGAGAAAATTGTCTGTATTAGCTCACCCCTTTTCCTAGCATTCCTCG 19373
QY 211 sPAsPLeuArgSerProIleArgAlaGluArgGluIleuLeuArgAsnLeuHisPhePhe 231
Db 19372 ATGATCTCCGGAGCCCTATTTCAGAGCTGAACGGGAAGAACTCCGAAACCTGCATTTTTC 19313
QY 231 rPPeGluTyrGly----- 235
Db 19312 CACCAAGATATGCGGTGATGCTCCTGAGCTGCTCAGCATCTGACGGGCTAGCACATC 19253
QY 236 -----GluValAlaAspGlyThrAlaVal 243
Db 19252 AGAAGTGAACCAAGCTGTCTTGTCTTCCCATTTCAAGAAATCAAGATGGAAGCTGTG 19193
QY 244 AspIlePheSerPheGlyMetCysValAlaGlu----- 254

Db 19192 GACATCTTCTCTTCGGAGATGTGTGCACTGGA-GGTACTGTCCACTCTCCATAGCCCC 19134
QY 255 -----MetAlaVal 257
Db 19133 TGCCTCAGGGGTCTAGCCATCCCAACATTCATTTGTCCCTGCACCGAGATGGCTGTA 19074
QY 258 LeuGluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArg 277
Db 19073 CTGAGATCCCAAGCCACCGGGATACCAAGATCAGAAAGAGGCATTCGCTGAGGCCAG 19014
QY 278 HisSerLeuSerAspProAsnMetArg----- 286
Db 19013 CACTCACTGATGACCCCAACATGGCGGTGAGAGACTTACTTGCCCTGTGGGCTGGC 18954
QY 286 ----- 286
Db 18953 TGCAGGTCTTAGCTTACCTTGAAGTCTTAGCAGGCGCAGAGGGGACGCTGTGC 18894
QY 286 ----- 286
Db 18893 GTAAATCTGAGAGGGGACGCTGTGTCACTAATCTGAGAGTGTAGCAGAGAGCTAAGGC 18834
QY 286 ----- 286
Db 18833 CCACTGAAGAAGACTCTTCCAGAAAGCTCAAGACTCCACAAAGGGTGAATGGAATAG 18774
QY 286 ----- 286
Db 18773 GATACGTGTGTGAGCCCAAGGCTTGCTGTCTCAAGTGTCCAGAAATCAGATCTTGAG 18714
QY 286 ----- 286
Db 18713 CTGCTAATGGATGACTTCTGTCTCATTCATACAGAAAGGCTTTTCAGATCCCTGAG 18654
QY 286 ----- 286
Db 18653 GTGGAGATGTGTGATTTGACCGGTGGGGTCTAGAGCTTGAAGAGACTCTCTAGAG 18594
QY 287 -----GluPheIleLeuCy 291
Db 18593 CGAGGACTTTTCAGCTATCACCTGTCTTCTGTTCACAACTGAAATTCATCTCTC 18534
QY 291 sCysLeuAlaArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgVal 311
Db 18533 CTGCTGGCCCGGAGCCCTGCCCGCAGACCTTCACCACTCTCTCTTCCAGGAGT 18474
QY 311 lleuPheGluValHisSerLeuIleLeuAlaHisCysPheIleGlnHisGln-T 331
Db 18473 GCTCTTTGAGTGACCTGCTGAAGCTGTGSCAAGCTACCTCTTCATTCAGACACAGTG 18414
QY 331 YF----- 331
Db 18413 TGAAGGAGAGAAAGGCGCGGGTGGGTGAGATCTGGGGTCAAGTATGTGAAGGCTTT 18354
QY 332 -----LeuMetP 334
Db 18353 GCAGGACCTTGAAGTGGGTCCATCCCATGAGCACTGTGTGTTTCACAGACTTCATATC 18294
QY 334 rGluAsnValAlaGluGluTyrThrLysAlaMetAspLeuHisAlaValIleuAlaGlu 354
Db 18293 CTGAGAAATGTGTAAAGAAAAGACCAAGGCCATGAACTCCATCAGATTTTGGCTGAGA 18234
QY 354 euProArgProArgArgProProLeuGlnTrpArg----- 365
Db 18233 TGCAGAGCCCATGAGACCCCAATGAGTGGCG- GTGAGTAGTGGCAACGAGGCTTC 18175
QY 365 ----- 365
Db 18174 CCTTCTAGAGCTGTGGCTTAGGCTTTGCCCTGACATCTCTCATGTACTTCACACCAT 18115
QY 366 -----TyrSerGluValSerPheMetGluLeuAspLysPheLeuGluAspVal----- 381
Db 18114 CTAGTACTCAGAGGTCTCTTCTTGGAGCTGGAACAAATTCCTAGAGAGATGT-CAGGTGA 18056

```

Oy 381 ----- 381
Db 18055 GAGCTGSCATGAGATCGAAGAAAAACACTATACAGCCCTGAGCCCCCACTAGTACCCA 17996
Oy 382 -----ArgAngrjY11eTyProleuWeeAsnPha1a1aThrArp 396
Db 17995 CCCCCTGTGTATCCCGAGAAACGGGATCTATCACTGATGAATTTGGGGCTGTCTGGC 17936
Oy 396 roLeuGlyLeuPrArgrValLeu1aProProProGluVal1Gln1ySal1AltyrPr 416
Db 17935 CTTGGGGCTTCCCGTGTGTGGCCCAACCCAGAGAAAGCCCAAAAGGCCAAACATC 17876
Oy 416 roThrProGluProPhaAspSerGluThrArgLyVal11e----- 429
Db 17875 CAAGCCCAAGAACCTTTGACTCGAGACCAAGAAAGGTGAG-TCCCACTCTGCTGTACT 17817
Oy 429 ----- 429
Db 17816 GTGCTGTCTTAGAGATTGACCAACAATGAGGAATGAAATGATGCTTTCTGTGGGA 17757
Oy 429 ----- 429
Db 17756 CCGGCTCTCCCATCCCGTTCTCCAGGCTGCCAGAGCAAAATCCAAGTGAACC 17697
Oy 429 ----- 429
Db 17696 CTTTGTGAGAGCTGGAACCGGACCAACAATCATTTCCCAAGAGCTTTCTTGTACT 17637
Oy 429 ----- 429
Db 17636 AGTCTTCCCATCCGCTGCTGTCTTCCAGAGGCTTTGGCTACGTCTGTCC 17577
Oy 429 ----- 429
Db 17576 GTCTGATCCAGAGCAATGTCTCCGCTCTCCCTTCTCAGAGCTTCTGTCTCAATA 17517
Oy 429 ----- 429
Db 17516 CACCTTAATTCATTCGCAATGCTCTCATGAGCTGCCAGAGAACTAGTGAATCAG 17457
Oy 429 ----- 429
Db 17456 CCTAAGCCCAAGCCTCTGTCTCTCTGCTGAGCACTTGTTTCTGTACAGTCAAC 17397
Oy 429 ----- 429
Db 17396 CCTGGGGGACAGCGTTCTTTGCTTCCAGGTTCTTTCTGTCTGTCTGTCTGC 17337
Oy 429 ----- 429
Db 17336 CTCCCCCATCCCTCTTGGCAATTGTATTATTCATTCAGAACCAACCAAGGCTGTCT 17277
Oy 429 ----- 429
Db 17276 TGAAGCTTCCCTAATACAGAGGCTGTGAGTAAGAAGTCTTTCCGACTTCCCTTG 17217
Oy 429 ----- 429
Db 17216 GAGACATTATACATCTTGTGGTCTGTGTCTCTTCTGAGGCTCTCTGTAAAGCCGTG 17157
Oy 430 -----GlnMeGlnHyAsnLeuGlu 437
Db 17156 CTCTGTGCTAATACACTTCCCTGCCCCAGAGTGTCTCAATGACAGTCAACCTGGAA 17097
Oy 437 TgSerGluAspLyAsn1aArgrTrpHis----- 445
Db 17096 GAAGGAGAGACAGGCTCGGTGGACGTGAGCAAGGGGCTGTGAGACTAGGAAT 17037
Oy 446 -----LeuThrLeuL 449
Db 17036 GGGGACCAATCAACTGTGGCGCGCTGTAGACACGCTTCCGTCTGAGCTTACTCTGC 16977

```

```

Oy 449 euleVal1LeuGluAspArgLeuHisArgGlnLeuThrTyAspLeuLeuProThr----- 467
Db 16976 TCTTGATCTTAGAGACCGGCTTACATCGCAGCTGACCTATGATCTGCTCCCAAG-TAGG 16918
Oy 467 ----- 467
Db 16917 CTGGATGAGGGGTGACAGGTGTGGGCTGAGGCTCAGGCACTGGCTAGTGCATGACTATT 16858
Oy 468 -----AspSer1aGlnAspLeu1aSerGluLeuValHisT 480
Db 16857 TGTCTGTGTCTTCTCTTCAAGCGGACAGTGCAGAGACTGTGCTGAATAGTGCATT 16798
Oy 480 YrGlyPheLeuHisGlu----- 485
Db 16797 ATGCTTCTCTCAGCA-GGTGGCGGGTGCCTTTGGGTGTGAAAGGTGCAAGGCTCTGC 16739
Oy 486 -----AspAspArgMetLysL 491
Db 16738 CGAGCTGCCGGTACTACATGCTTCCCTACTCATCCAGTCAAGATGACAGAACAAAGC 16679
Oy 491 eul1a1aPheLeuGluSerThrPheLeuLyTyArgGlyThrGln1a 507
Db 16678 TAGCAGCTTCTTGAGACCACTTTCTCAAGTACCGAGGACCGAACCG 16629

RESULT 11
AC051615 213817 bp DNA linear HTG 11-SEP-2002
LOCUS DEFINITION Mus musculus strain C57BL6/J clone RP23-208H22, WORKING DRAFT
SEQUENCE, 46 unordered pieces.
AC051615
AC051615.10 GI:22779498
HTG: HTGS PHASE1; HTGS_DRAFT.
MUS musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
1 (bases 1 to 213817)
Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shm,C., Decker,J., Thomas,E., Pereira,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
High Throughput Mouse Sequencing
Unpublished
2 (bases 1 to 213817)
Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shm,C., Decker,J., Thomas,E., Pereira,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
Direct Submission
Submitted (15-APR-2000) Department of Molecular Genetics, Albert
Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
Bronx, NY 10461, USA
3 (bases 1 to 213817)
Han,J., Montgomery,K.T., Grills,G., Lee,E., Long,J., Pomerantz,R.,
Ioshikhes,I.P., Shm,C., Decker,J., Thomas,E., Pereira,A.,
Gordon,M., Goltz,J.S. and Kucherlapati,R.
Direct Submission
Submitted (11-SEP-2002) Harvard Partners Center for Genetics and
Genomics, Harvard Medical School, 65 Landsdowne St, Cambridge, MA
02139, USA
COMMENT
-----Genome Center
Center: Harvard Partners Genome Center
Center Code: HPGC
Web site: http://www.hpcg.org/Sequence/mouse.html
Contact: hpcg@medel.mgh.harvard.edu
-----Summary Statistics
Center project name: ACC
Sequencing vector: pUC18, L08752
Chemistry: Dye-terminator Big Dye, 100%
*Consensus quality: 207678 at least Q20
*Consensus quality: 204383 at least Q20
*Consensus quality: 198707 at least Q40
Estimated insert size: agarose-FP - N/A

```

**Estimated insert size: 212917 - sum-of-contigs
Quality coverage: agarose-FP - N/A
Quality coverage: 6.8 x in Q20 bases; sum-of-contigs estimation

NOTE: This is a 'working draft' sequence. It currently consists of 46 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 71517: contig of 71517 bp in length
* 71518 71537: gap of unknown length
* 71538 117192: contig of 45655 bp in length
* 117193 117212: gap of unknown length
* 117213 133385: contig of 16173 bp in length
* 133386 133405: gap of unknown length
* 133406 154137: contig of 20732 bp in length
* 154138 154157: gap of unknown length
* 154158 163981: contig of 9824 bp in length
* 163982 164001: gap of unknown length
* 164002 169740: contig of 5739 bp in length
* 169741 169760: gap of unknown length
* 169761 173809: contig of 4049 bp in length
* 173810 173829: gap of unknown length
* 173830 176074: contig of 2245 bp in length
* 176075 176094: gap of unknown length
* 176095 178772: contig of 2678 bp in length
* 178773 178792: gap of unknown length
* 178793 181465: contig of 2673 bp in length
* 181466 181485: gap of unknown length
* 181486 182101: contig of 616 bp in length
* 182102 182121: gap of unknown length
* 182122 184693: contig of 2572 bp in length
* 184694 184713: gap of unknown length
* 184714 185952: contig of 1239 bp in length
* 185953 185972: gap of unknown length
* 185973 188487: contig of 2515 bp in length
* 188488 188507: gap of unknown length
* 188508 189234: contig of 727 bp in length
* 189235 189254: gap of unknown length
* 189255 190461: contig of 1207 bp in length
* 190462 190481: gap of unknown length
* 190482 190657: contig of 186 bp in length
* 190658 190687: gap of unknown length
* 190688 192310: contig of 1623 bp in length
* 192311 192330: gap of unknown length
* 192331 194183: contig of 1853 bp in length
* 194184 194203: gap of unknown length
* 194204 195637: contig of 1434 bp in length
* 195638 195657: gap of unknown length
* 195658 196390: contig of 733 bp in length
* 196391 196410: gap of unknown length
* 196411 196937: contig of 527 bp in length
* 196938 196957: gap of unknown length
* 196958 198075: contig of 1118 bp in length
* 198076 198095: gap of unknown length
* 198096 199358: contig of 1263 bp in length
* 199359 199378: gap of unknown length
* 199379 200436: contig of 1058 bp in length
* 200437 200456: gap of unknown length
* 200457 201048: contig of 592 bp in length
* 201049 201068: gap of unknown length
* 201069 201765: contig of 697 bp in length
* 201766 201785: gap of unknown length
* 201786 202164: contig of 379 bp in length
* 202165 202384: gap of unknown length
* 202385 202997: contig of 813 bp in length
* 202998 203017: gap of unknown length
* 203018 204899: contig of 1882 bp in length
* 204900 204919: gap of unknown length
* 204920 205365: contig of 446 bp in length

```

FEATURES

source

```

* 205366 205385: gap of unknown length
* 205386 205759: contig of 374 bp in length
* 205760 205779: gap of unknown length
* 205780 205953: contig of 174 bp in length
* 205954 205973: gap of unknown length
* 205974 206484: contig of 511 bp in length
* 206485 206504: gap of unknown length
* 206505 206764: contig of 260 bp in length
* 206765 206784: gap of unknown length
* 206785 207831: contig of 1047 bp in length
* 207832 207851: gap of unknown length
* 207852 208358: contig of 507 bp in length
* 208359 208378: gap of unknown length
* 208379 209487: contig of 1109 bp in length
* 209488 209507: gap of unknown length
* 209508 210063: contig of 556 bp in length
* 210064 210083: gap of unknown length
* 210084 210488: contig of 405 bp in length
* 210489 210508: gap of unknown length
* 210509 211550: contig of 1042 bp in length
* 211551 211570: gap of unknown length
* 211571 212106: contig of 536 bp in length
* 212107 212126: gap of unknown length
* 212127 212454: contig of 328 bp in length
* 212455 212474: gap of unknown length
* 212475 212763: contig of 269 bp in length
* 212764 212763: gap of unknown length
* 212764 213236: contig of 473 bp in length
* 213237 213256: gap of unknown length
* 213257 213817: contig of 561 bp in length.

Location/Qualifiers
1..213817
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP21-208H22"
/sex="male"
1..71517
/note="assembly_name:Contig205"
71518..71537
/estimated_length=unknown
71538..117192
/note="assembly_name:Contig199"
117193..117212
/estimated_length=unknown
117213..133385
/note="assembly_name:Contig198"
133386..133405
/estimated_length=unknown
133406..154137
/note="assembly_name:Contig197"
154138..154157
/estimated_length=unknown
154158..163981
/note="assembly_name:Contig196"
163982..164001
/estimated_length=unknown
164002..169740
/note="assembly_name:Contig195"
169741..169760
/estimated_length=unknown
169761..173809
/note="assembly_name:Contig194"
173810..173829
/estimated_length=unknown
173830..176074
/note="assembly_name:Contig193"
176075..176094
/estimated_length=unknown
176095..178772
/note="assembly_name:Contig191"
178773..178792

```


QY 255 -----Met1Val 257
Db 14768 TGCTCAGGGGTCTAGCCCATCCACCAATTTGTCCTCGACCCAGATGGCTGTA 14827
QY 258 LeuGluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArg 277
Db 14828 CTCGAGATCCAAAGCCAAAGGGGATACCAAGATCACAGAAAGGCCATCGCTGAGCCAG 14887
QY 278 HisSerLeuSerAspProAsnMetArg----- 286
Db 14888 CACTCACTAGTGAACCCCAACATGCGGTGAGCAGCTTACCTTGCCCTCGGGGCTGAC 14947
QY 286 ----- 286
Db 14948 TGCAGGTGTTGACGTTACCTTTGAAGTGTCTAGGACAGGCCAGAGGGGACGCTGTG 15007
QY 286 ----- 286
Db 15008 GTAAATCTGAGAGGGGACGCTGTGTCAGTACATCTGAGAGTGTAGCAGAGGACTAAAGC 15067
QY 286 ----- 286
Db 15068 CCACTGAAGAAAGACTCTTCAGAACTCAGACTCAGAGGGGTAAATGAAATGAG 15127
QY 286 ----- 286
Db 15128 GCATACGTGTTGGAGCCCAAGGCTTGCTGTCTCAAGTGTCCAGATCAATCTTGAG 15187
QY 286 ----- 286
Db 15188 CTGTAATGGAGTGAATCTGTCTGTCAATTCAATACAGAAAGGCTTTTACATCCCTGAG 15247
QY 286 ----- 286
Db 15248 GTGGAGATGTGTGATGTGAACCGGTGGGTCTAGAGTCTTGAAGAACCTCTCTAGAG 15307
QY 287 -----GluPheIleLeuCy 291
Db 15308 CGAGACTTTTCCAGCTATCACTGTCTTTCTTTTACACCTTGAAGTATATCTCTC 15367
QY 291 sCyLeuAlaArgAspProAlaArgAspProSerAlaHisSerLeuLeuPheHisArgVal 311
Db 15368 CTGGCTGGGCGGGAGACCTGGCCGCGAGACCTCAGGCCCAACCTCTCTTCCACGAGT 15427
QY 311 LeuPheGluValHisSerLeuLeuLeuAlaAlaHisCyAspPheIleGlnHisGln-T 331
Db 15428 GCTCTTTGAGTGTGACTCGCTGAGCTGTGGCAGCTCTGCTTATCATCAGACACAGTG 15487
QY 331 YF----- 331
Db 15488 TGAGGGGCAAGAGCGCGGGGTGGGTCAAGATCTGGGGTCAAGTATGGAAGGTCTTT 15547
QY 332 -----LeuMetP 334
Db 15548 GCAGGACCTTAGTGGTGGTCCATCCCATAGACATTGTGTCTTACACAGACCTCATGC 15607
QY 334 TGGluAsnValValGluGluValThrLysAlaMetAspLeuHisAlaValLeuAlaGluL 354
Db 15608 CTGAAATATGTGTAGAGAAAGAACCAAGGCCATGACCTCATGCAAGTTTGGCTGAGA 15667
QY 354 euProArgProArgArgProPheLeuGlnTPrArg----- 365
Db 15668 TGCCTGACGCCATGAGACCCCAATGACAGTGGCG-GTGAATAGTGGCCAAAGGAGCTTC 15726
QY 365 ----- 365
Db 15727 CCTTCTAGCTGCTTGGCTAGGCTTTGCCCTGACATCTCTCATGTATCCTTCCACACAT 15786
QY 366 -----TyrSerGluValSerPheMetGluLeuAspLysPheLeuGluValPVal----- 381
Db 15787 CTAGTACTCAGAGGCTCTCTCTTCTTGAGCTGGAACAATTTCTAAGAGATG-CAGGTGA 15845
QY 381 ----- 381

Db 15846 GAGCTGCATGAGACTGAAGAAAAAGCACTATCAGGCTGAGCCCACTAGCTAACCA 15905
QY 382 -----ArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgP 396
Db 15906 CCCCTTGTCTATCCCAAGAAACGGGATCTATCCTATGAACTTTGGGCTGCTGGC 15965
QY 396 roLeuGlyLeuProArgValLeuAlaProProGluGluValGlnLysAlaLysThrP 416
Db 15966 CCTTGGGCTTCCCGTGTGTGGGCCCAACCCCAAGAGAACCCAAAGGCCAAACATC 16025
QY 416 roThrProGluProPheAspSerGluThrArgLysValIle----- 429
Db 16026 CAACGCCAGAACCTTTGACTCGAGAACCAAGAGGTGAG-TCCCACTCTGCTGTACT 16084
QY 429 ----- 429
Db 16085 GTGTGTCTCTAGAGATTGACCACAATCAGGGATGAAATGATGATGCTTTCTGGGGA 16144
QY 429 ----- 429
Db 16145 CCGGCTCTCCCATCCCGTTCCTCCAGGCTGCAGAGACAAATCCAAGCTGACCC 16204
QY 429 ----- 429
Db 16205 CTTTGTGAAGCTGAAACCGAACCGACACATCATTTCCCAAGGCTTCTTGTTC 16264
QY 429 ----- 429
Db 16265 AGTCTTCCCATCCGGTGTGATGCTCTCTTCCAGAGGCTTTGGCTAGTGTCTTCCC 16324
QY 429 ----- 429
Db 16325 GTCTGATCCAGAGCAATGTCTCCGCTCTCTCTCAATAGCTTCTGTCTCAATA 16384
QY 429 ----- 429
Db 16385 CACCTAATTCATTCAGTGAATGTACATGAGCTGTCCAGAGAACTAGTGAATAACG 16444
QY 429 ----- 429
Db 16445 CTTAAGCCCAAGCTCTGTCTCTCTCTGTGTGACCTGTGTTTCTGTACAGTCACAC 16504
QY 429 ----- 429
Db 16505 CCTGGGGGCAAGCGTTTCTTGTCTTCAAGGTTCTCTTCTGTCTGTCTGTG 16564
QY 429 ----- 429
Db 16565 CTCCCCCATCCCTCTGTGGCATTTGATTCATTCAGAACCCAAACGAGCTGTCT 16624
QY 429 ----- 429
Db 16625 TGAAGCTTCCCTAATACTCAGAGGCTGTGACGTAAAGCTGTTTCCGACTTCCCTTG 16684
QY 429 ----- 429
Db 16685 GAGACATTATCACTTTTGGGTCTGTGTCTCTCTGTGGGCTCTCTGTAAAGCCGTG 16744
QY 430 -----GluMetGlnCyAsnLeuGluVal 437
Db 16745 CTGTGTGCTAATACCTTCCCTGCGCACAGGTGTCTCAGATGTGCAACCTGAAA 16804
QY 437 rGSerGluAspLysValAlaArgTPrHis----- 445
Db 16805 GAAAGGAGAGCAAGCTCGGTGTGACGTGAGCAAGAGGCGCTGTGAGACTAGGAAAGAT 16864
QY 446 -----LeuThrLeuL 449
Db 16865 GGGGACCAATCACTGTGGCGCGCTGTGACACAGCCCTTCGCTGTGACGCTTACTGTC 16924
QY 449 euLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr----- 467

```

Db      16925 TCTGTGCTTGTAGAGACCGGCTACATCGGAGCTGACCTATGATCTGCTCCCAAG-TAGG 16983
Oy      467 -----
Db      16984 CTGTAGTGGGGTCAAGTGTGGCTGAGGCTCAGGCACTGGCTAGTGCATGACTCATT 17043
Oy      468 -----
Db      17044 TGTCTGTGTTCTCTCTTTCAGGAGGACAGTCCCGACGACTCTGCTCTGAACTAGTGCAATT 17103
Oy      480 YTGlyPheLeuH1g1u-----
Db      17104 ATGGCTTCTGACACA-GATGGCGGGGTGGCGCTTTGGGTTGGAAGGTGACAGGCTTGC 17162
Oy      486 -----
Db      17163 CGAGTGGCCGGTACTACATGCTTCCCTACTCATCATCAGTGCAGGATATCAGGACAAAGC 17222
Oy      491 eua1a1aPheLeuH1g1uSerThrPheLeuYrYrYrG1YThrG1na1a 507
Db      17223 TAGCAGCTTTTGGAGACACTTTTCTCAGTACCGAGGAGCGAAGCG 17272

RESULT 12
AC132640 194709 bp DNA linear HTG 15-NOV-2002
LOCUS AC132640.3 GI:25007393
DEFINITION Rattus norvegicus clone CH230-339B14, WORKING DRAFT SEQUENCE.
ACCESSION AC132640.3
VERSION AC132640.3
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
          Sciurognathi; Muridae; Muridae; Murinae; Rattus.
REFERENCE
  1 (bases 1 to 194709)
  Murny,D,Marie., Metzger,M,lee., Abramson,S., Adams,C., Alder,J.,
  Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
  Ayaalbechi,Y., Ayagci,A., Ayodeji,M., Baca,E., Baden,H.,
  Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
  Biwalto,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
  Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
  Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
  Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
  Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
  Davila,M.U., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
  Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Ditya,K.,
  Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
  Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
  Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
  Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
  Georgegeorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
  Gunaratne,P., Haaland,W., Hamli,C., Hamilton,C., Hamilton,K.,
  Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
  Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
  Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
  Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Joliver,A.,
  Kapathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
  Kowals,C., Kraft,C.T., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
  Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
  Lorenshaw,L., Louisedge,H., Lozada,R.J., Lu,X., Ma,J.,
  Maheshwari,M., Mahindartine,M., Mahmoud,M., Malloy,K., Mangum,A.,
  Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,B.,
  Mawhiley,S., McLeod,M.P., McNeill,T.Z., Meenen,B.,
  Miliavajevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
  Morgan,M., Morris,K., Morris,S., Mundasa,M., Murphy,M., Nair,L.,
  Narkervic,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
  Nwokilemeh,O., Okwomu,G., Olampunagoon,A., Pal,S., Parks,K.,
  Patelrnak,S., Paul,H., Perez,A., Perez,L., Pflankoch,C.,
  Plopper,F., Poindexter,A., Popovic,D., Prins,B., Pu,L.-L.,
  Puzo,B., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
  Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
  Rives,C., Rokeby,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
  Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Gendank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 194709: contig of 194709 bp in length.
Location/Qualifiers
1. 194709
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-339B14"
1. .3065
misc_feature

```


Dh 53564 GCACATCAGAAATGTCATCAACTGCTGTTTTCATTCACAGAGTCACAGATGG 53505
Qy 240 YTHAlaValAlapIlePheSerPheGlyMetCysAlaLeuGlu----- 254
Db 53504 GACTGCTGTGACATCTTCTCTCTTGGATGTGTGACATGGA-GGTACTGTTCACTCTCC 53446
Qy 254 ----- 254
Db 53445 TCATAGCCTCTGCTTGGCTCTACTCATCTCACCATTTTTCATTTGTCCTGCACTCA 53386
Qy 255 -MetAlaValLeuGluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAl 274
Db 53385 GATGCGGCTCTTGGAGATCCAAAGCGGGAATACCGAGTCCAGAAAGAGCTATTGC 53326
Qy 274 AATGAlaArgHisSerLeuSerAspProAsnMetArg----- 286
Db 53325 TCGGGCCAGGACCTCGCTGAGTGAACCCACATGCGGTGAGACAGCTTAACCCCTTAGG 53266
Qy 286 ----- 286
Db 53265 ACTGCTGCAAGTCTTGGGCTTGCCTTGAAGCCCTTAGTAGGCTAGAGGTATGTTCT 53206
Qy 286 ----- 286
Db 53205 GGTCAATCATTTGAAGTGTATGATCTGAAGCGTAGCTGGGGACTGAGTCACTC 53146
Qy 286 ----- 286
Db 53145 AAAAGAAAGCTCTCTAGAGCTCACGGAATTTCAAGAGGTGAATGAGTAGGGCATG 53086
Qy 286 ----- 286
Db 53085 GGTGTTGGAGCCCAAGGCTTCTGCGCTCAAGTTTACAGAAATCAGATCCTGAGCTCT 53026
Qy 286 ----- 286
Db 53025 AATGGATGCTTCTCAGGACGAGCATGATGAGAGCATATTGGAGGTGGAGAGATGC 52966
Qy 286 ----- 286
Db 52965 AATGATTGACCCGTGGGCTTAGGGTCTTGAAGATCTCTTAGACGAGATTTCCTC 52906
Qy 287 ----- 287
Db 52905 ACCCACCACTACTTCTTCCCAACACTTAGAAATTCATCTCTCTGCTGCGCGGG 52846
Qy 296 SPProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluAlaH 316
Db 52845 ACCGTGCGCGCGACCTCAGCGCCCAACCTCTCTCCACCGAGTGTCTTTGAGGTGC 52786
Qy 316 HisSerLeuLeuLeuAlaAlaHisCysPheIleGlnHisGln-Tyr----- 331
Db 52785 ACTGCTGAGAGCTGTGACAGCCCACTGTTTATCATCAGACACAGTGTAGGGGCGCAGAGG 52726
Qy 331 ----- 331
Db 52725 GCCGGGTGACACAGGATCTGGGGTCAATCATGAAAGGCTTTTGACGAGACCTTAAG 52666
Qy 332 ----- 332
Db 52665 TTGGGTCATCCCATGCGGCAATTTCTGTGTTTCAACAGACTCATGCTCGTGAAGATGTGTA 52606
Qy 339 GluGluValThrIleValMetAspLeuHisAlaValLeuAlaGluLeuProArgProArg 358
Db 52605 GAGGAAAAGACCAAGGCACTGACCTCATGACAGTGTGCTGAGATGCCCAAGCCCAT 52546
Qy 359 ArgProLeuGlnTyrArg----- 365
Db 52545 GGAACCCCAATGCAATGAGCG-GTGAAGTGTGCAAAATAGAGCTTCCCTCTAACTGCTT 52487
Qy 365 ----- 365
Db 52486 GGCTAGGCTTTGGCTGTATGCTCTTATGTATGACACCTTTCATACATCCAGG 52427

Qy 366 TyrSerGluValSerPheMetGluLeuAspIlyPheLeuGluAlaVal----- 381
Db 52426 TACTCAGAGTTCCTCTTGGAGCTGGAACAAATTCCTTAGAGATGTCAGGTGAGAGCTG 52367
Qy 381 ----- 381
Db 52366 GCATGACAGTGAAGAAAAACATATCAGACCCGAGTCCCACTTACCTACCCCTGTCT 52307
Qy 382 -----ArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGlyLeu 399
Db 52306 ATCTCAGGAATGGCATCTATCCACTCATGAATTTGCTGTGCTCGGCTTGGGAATT 52247
Qy 400 ProArgValLeuAlaProProProGluGluValGlnIlyAlaIlyThrProThrProGlu 419
Db 52246 CCTGTGTGTGGCCCAACCCCAAGAGGCGCCAAAAGGCCAAATCCAAACCCCAAGAA 52187
Qy 420 ProPheAspSerGluThrArgIlyS----- 427
Db 52186 CCTTTGACTCGAGACCAAGAGGTAGTCCCAATCTGCTGTACTATGCTGTTTCTC 52127
Qy 427 ----- 427
Db 52126 AGAGTTAACACACACAGGAAATGGTGCCTTGGGACCTGTCTGTCCATCCAAAT 52067
Qy 427 ----- 427
Db 52066 CTCCCGAGCTTCCAGAGCAAAATCCAAAGTGAACATTTTCCAGAGGCTTTTGC 52007
Qy 427 ----- 427
Db 52006 TTTTACAGTCTTCCCATCCATGCGTGTCTCTTCATAGGCTTTTGGCACTTCA 51947
Qy 427 ----- 427
Db 51946 TTCCCATGTGATCCAGAGCATCACTGTGATCTCCCACTCTAGCTTCATCTTCC 51887
Qy 427 ----- 427
Db 51886 TCAGCACACTAAATTCAAATGCAATCGTATCATGATCTGTGATCCAGCTGACAGA 51827
Qy 427 ----- 427
Db 51826 GAGCTAGTAAGAAATCAGCTCAGGGGTTGGGATTTAGCTCAGTGTAGAGCGCTTGC 51767
Qy 427 ----- 427
Db 51766 TAGGAAGCTAAGGCTGTGGTGGTCCCACTCCGAAAAAAGAACCAAAAAA 51707
Qy 427 ----- 427
Db 51706 AAAAAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAATCAGCTCAGTCCCTGT 51647
Qy 427 ----- 427
Db 51646 CTTGCTCTGTGACACCCATTTTCTGTACTGTGACCCCTGAGACACAGCATTTCT 51587
Qy 427 ----- 427
Db 51586 TGCTTTCACAGATTCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 51527
Qy 427 ----- 427
Db 51526 GTCTGTCTCTCTGTGTGTGTGTGTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTG 51467
Qy 427 ----- 427
Db 51466 CATTGTATTCATCATTCAGTACCAACAGAAAGCTACTACTTGTGAAGCTTCCCTAAT 51407
Qy 427 ----- 427
Db 51406 ACTCAGGATGTGTCAGTAAAGATGTGTTTCCGGACTTCCCTTGAGAGATGTCTACAT 51347


```

Qy 427 ----- 427
Db 51346 TTTTGGGCTCTGCTGATCCTTCTGGGCTCTTTGTTAATAATACCTGTGGCTATAC 51287
Qy 428 ----- Val11g1mctg1ncysaantleu1uargsegl1uaplyal1 442
Db 51286 ACTTCCCTGCCACAGCTGTGTCAATGACATGCAACTGGAAAGAGGAGCAAGAGC 51227
Qy 442 AATGTTPHIS----- 445
Db 51226 TGGGTGGACAGTGACGAAGGGCCCTGTGGGAGCTGGAAGAAAGCGGGACCAATCACTC 51167
Qy 446 ----- Leuth1leu1leu1Val1leu1lubs 454
Db 51166 TGGACACCGGCTGACCAAGCCCTTCCGTGACAGCTTACTCTTTGGTGTGAGGA 51107
Qy 454 PARGLEUHIARGINLEUTH1TYRASP1LEU1PO----- 466
Db 51106 CCGGTATACATCGGACAGCTGACCTATGATCTGCTCCAGATAGGCTGTAGGGAGATCAGGT 51047
Qy 466 ----- 466
Db 51046 GGTGGGCTGGGGCTACAGGACCGGCTAGTGCCAACTGACTTGTCTGTCTCTT 50987
Qy 467 -Thr-AspSer1agi1napleu1aser1leu1Val1st1yrg1yPhe1u1ag1u- 485
Db 50986 CAGCGACAGCTGCCAGACGACTCTGCGCGGAACTAGTGCATGTGCTTCTGACAGG 50927
Qy 485 ----- 485
Db 50926 TGGGCGGGTGGCGCTGCGGTTTGGAAAGGTGACGTGCGGAGCTGCCAGTACTACCGT 50867
Qy 486 ----- Aspa1arg1me1lyleu1al1a1phe1u1g1user 497
Db 50866 GCTTCCCAACCCCTTCCACTGACGAGTACCGGACCAAGCTCTTCTGAGACC 50807
Qy 498 ThrPhe1u1y1e1y1r1g1y1Thrg1na1a 507
Db 50806 ACTTTTCTCAAGTACCGTGGACACAGCG 50777

RESULT 13
AC115184 248568 bp DNA linear HTG 19-NOV-2002
LOCUS AC115184
DEFINITION Rattus norvegicus clone CH230-52N10, WORKING DRAFT SEQUENCE, 4
unnumbered pieces.
ACCESSION AC115184
VERSION AC115184.4 GI:24635546
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Murioidea; Muridae; Murinae; Rattus.
1 (bases 1 to 248568)
Muzny D,Marie, Metzker M, Lee, S., Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D.,
Aryalasech, V., Ayagi, A., Ayodeji, M., Baca, B., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Barnham, F.,
Biswal, K., Blais, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Cesari, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Ercotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebremeseg, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hognes, M.,

```

REFERENCE

Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshew, L., Louised, H., Lozada, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackeleme, O., Okwomou, G., Olarnunagsoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Woodem, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, J., Zhao, S., Zhao, S., Dunn, D., von Weinstock, G., and Gibbs, R.A.

Direct Submission
Unpublished
2 (bases 1 to 248568)
Worley, K.C.
Direct Submission
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 248568)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (19-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 6, 2002 this sequence version replaced gi:23267220.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSCY
Center clone name: CH230-52N10
----- Summary Statistics
Assembly program: phrap, version 0.990329
Consensus quality: 230402 bases at least Q40
Consensus quality: 233092 bases at least Q30
Consensus quality: 234313 bases at least Q20
Besticmed insert size: 237322; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

Dh	86874	CAGGTGAGAGCCTTGCTTGCCCAAGCACTGGCCCTTCTCTTCCTTCACTTTCTTTCC	86933
Qy	161	-----Leu-SerPheLeuHia	166
Dh	86934	GGTGGCTCTGACCCGGTTTTTCTCGCTCCCTTCCTCTAGTGTGAAGTTTTTTCAGG	86993
Qy	166	laCySerProProIleIleHieglYanLeuThrSerApThrIlePheIleGlnHia	186
Dh	86994	CCTGCAGTCCCCCATCATTCACCGAAACCTGCACAGCGACAGATCTTCATTCAACACACA	87053
Qy	186	angIYleuHieglYSer-----	193
Dh	87054	ATGGCTTCATCAAGATCGCTCTGTGCTGGCAGCGGGCGAGGGGACAGGGGT	87113
Qy	193	-----	193
Dh	87114	GGGGAACAGGGGAGGCCAGAGACAGACTGCCAGTCTCTGGCCCCACTGACGCT	87173
Qy	194	-----ValTrpHieArgIlePheSerAn-----	201
Dh	87174	CTCTTTCTCTGCGTGGCCGTGTGACAGTGTGNACCGACTCTTCTCAATGGTGTGG	87233
Qy	202	-----AlaLeuArgProPothr	207
Dh	87234	GCTCTCGGGGTGGAGCTGTGGAAACTGGTCTGCTCAACATCACATTTTCTTA	87293
Qy	208	AlaLeuProApApLeuArgSerProIleArgAlaGluArgGluIleuArgAnLeu	227
Dh	87294	GCACTTCTGTACGATCTCCGAGCCCTTAAGAATGAAACGGAGAGCTCCGAAACTTG	87353
Qy	228	HiAphePheProProGluTrgIY-----	235
Dh	87354	CACTTTTCACACAGATGATGGCGGTAGTCTTAAGCCGTTCCATCTGACCTGCTG	87413
Qy	236	-----GluValAlaApGcl	240
Dh	87414	GCACATAGAAATGCCAATCAACTGCTGTTTTCATTTCCATTCAGAGATCAACGATGG	87473
Qy	240	YThrAlaValaApIlePheSerPheGlyMetCysAlaLeuGlu-----	254
Dh	87474	GACTGCTGTGACATCTTCTCTTTGGAGTGTGTGACTGGA-GGTACTGTTCACTCTCC	87532
Qy	254	-----	254
Dh	87533	TCATAGCCTCTGCTTGGCTCTACTCATCTCAACATTCATTTGTGCCCTGCACCA	87592
Qy	255	-MetAlaValleuGluIleGlnThrangIyaApThrArgValTh-GluGluAlaIleAl	274
Dh	87593	GATGGCGGTCTTGGAGATCCAAAGCAACGGGGATACCCGAGTCAAGAAAGATATTGC	87652
Qy	274	ArgAlaArgHieSerLeuSerApProAnMetArg-----	286
Dh	87653	TGGGCGCAGGACTCGCTGAGTGACCCAAACATGCGGGTGAGCAGCTTAACCCCTCTAGG	87712
Qy	286	-----	286
Dh	87713	ACTGCGTCAAGGTCTTGGGCTTGCCCTTGAAGCCCTTAGTAGCTAAGGTATAGTCT	87772
Qy	286	-----	286
Dh	87773	GGTCAGTACATTGMAAGTGTAGACTGTGAAGCGTAGCTGCGAGTCACTC	87832
Qy	286	-----	286
Dh	87833	AAAGGAAGACTCTTTAGAGACTCAGGAATTTCAAGAGGGTGAATGAAATAGGGCATG	87892
Qy	286	-----	286
Dh	87893	GGTGTGAGCCCAAGGCTTCCTGCGCTCAAGTTTACAGATCAGATCTGAGCTGCT	87952
Qy	286	-----	286
Dh	87953	AATGGATGCTTCTCAGCAGGACATGATGAGACATCATTTGTGAGGTGAGAGATGC	88012

QY	286	-----	286
Db	88013	AATGATTGGACCCGTGGGCTCTAGGGCTTTGAAAATCTCTCTAGACGAGGATTTTCC	880172
QY	287	-----glupheilleucysCybLeu1aArga	296
Db	88073	ACCCACCACTACTTTCTTCTCCCAACCACTAGAAATTCACTCTCTCTGCTGGCCGGG	881327
QY	296	bpPro1aArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValH	316
Db	88133	ACCTGCCCCGACCCCTCAGCCCAACAACCTCTCTTCCACCAAGTCTCTTTGAGGTGC	881927
QY	316	IsSerLeuLybLeuLeuAlaAlaHisCybPhe1leg1nHsg1n-Tyr-----	331
Db	88193	ACTGCTGAGACTCTGGACGCCCACTGTTTATTCACGACCAAGTGAGGGGACAGAAG	882547
QY	331	-----	331
Db	88253	GCCGGGGGTGACAGATCTTGGGGTCAGATCATGMAAGGGCTCTTGACGACCTTTAG	883127
QY	332	-----LeuMetProGluMetVal1Val	338
Db	88313	TTGGGTCCATCCCATGGGCAATTTCTGTGTTTCAAGACTCTCAATGCTTGAATGTGGTA	883727
QY	339	GluGluArgThrLybAlaMetAspLeuHisAlaValLeuAlaGluLeuProArgProArg	358
Db	88373	GAGGAAAAAGACMAAGGCATGGAACCTTCAGCACTGTGTGGCTGAGATGGCCACGCCCAT	884337
QY	359	ArgProProLeuGlnTTPArg-----	365
Db	88433	GGAACCCCAATGACGTGGCG-GTGAGTGTGGCAAAAGAGCTTCCCTCTAACTGCTT	884937
QY	365	-----	365
Db	88492	GGCTAAGCTTTGGCTTATGCTCTTATGTCTTATGACACTTTCATACCATCCCAAG	885517
QY	366	TyrSerGluValSerPheMetGluLeuAspLybPheLeuGluAspVal	381
Db	88552	TACTCAAGAGGTTTCTCTTCTTGAGCTGGAACAATTCCTAGAGATGTCAAGTGAGAGCTG	886117
QY	381	-----	381
Db	88612	GCATGACAGTGAGMAAAAACTACTACAGCCGAGTCCCCCACTAACCCCCCTGTCT	886717
QY	382	-----ArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGlyLeu	399
Db	88672	ATCTCCAGGAATGGCATCTATTCACATCAATGAATTTGGCTGTGCTGGCCCTTGGGACTT	887317
QY	400	ProArgValLeuAlaProProProGluGluValGlnLybAlaLybThrProThrProGlu	419
Db	88732	CCTGTGTGTGGGCCCAACCCCAAGAGAGGCCCAAAAGGCCAAACTCAACCCACAGAA	887917
QY	420	ProPheAsnSerGluThrArgLybS	427
Db	88792	CCCTTTACTGTGGAGACCAAGAAAGTGAGTCCCACTCCTGCTGTACTATGCTGTTTCTC	888517
QY	427	-----	427
Db	88852	AGAGTTAACACAAACAGGGAATGGTGCTCTGGGGAACGTGCTGTCTGCATCCCAATT	889117
QY	427	-----	427
Db	88912	CTCCCCAGGCTTCAGAGACCAAAATCAAGCTGACCAATTTTCCCAAGAGGCTTCTTGC	889717
QY	427	-----	427
Db	88972	TTTACAGACTCTTCCCATCCCATGCGGTGCTCTTCCATFAGGCTTTGGCACCTGCA	890317
QY	427	-----	427
Db	89032	TTTCCCATCTGATTCAGAGCCATCACTGTATCTTCCCATCTAGCTTCCATCTTCC	890917

```

Oy 427 ----- 427
Db 89092 TCAGACACACTAAATTCAATGCGATGCTCATGTCTCGATCCAGCTGCACAGA 89151
Oy 427 ----- 427
Db 89152 GAGCTAGTAAATCAGCCTCAGGGGATTGAGGATTTAGCTCAGTGTAGAGCGCTTGC 89211
Oy 427 ----- 427
Db 89212 TAGAAGGCTAAGCCCTGAGTGGGTCCAGCTCCGAAAAAAGACCAGAAAAA 89271
Oy 427 ----- 427
Db 89272 AAAAAAAAAAGAAAAAGAAAAAGAAAAAGAAATCAGCTCAGTCCCTGT 89331
Oy 427 ----- 427
Db 89332 COTGCTCTGGTGGACCCCAATTTTCTGTACTGTGTGACCCCTGAGACAGCAATTCCT 89391
Oy 427 ----- 427
Db 89392 TGCTTCCACAGTCTCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTCTT 89451
Oy 427 ----- 427
Db 89452 GTCTGTGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 89511
Oy 427 ----- 427
Db 89512 CATTGTATTCACTTCACTAGTACCAGAAAGAGCTACTACTTGGAGCCTTCCCTAAT 89571
Oy 427 ----- 427
Db 89572 ACTTCAGGAGATGTGTCAGTAAAGATGTGTTCCGACTTCCCTGAGATGTACCAAT 89631
Oy 427 ----- 427
Db 89632 TTTTGGGTCTGCTGATCCTTCTGGGCTCCCTTTGTAATTAATCTCTGTGGCTAATAC 89691
Oy 428 ----- 442
Db 89692 ACTTCCCTGCACAGTGTGTCTCATGTGATGACCTTACTGCTTGTGTGTGTGAGGA 89751
Oy 442 AATGTTPHIS----- 445
Db 89752 TCGGTGACATGAGCAAGGCGCTGGGTGGGACTAGAGAAAGACGGAGCAATCACTC 89811
Oy 446 ----- 454
Db 89812 TGGCACCGCCCTGACACACGCCCCCTTCCGTGCACAGCTTACTGCTTGTGTGTGAGGA 89871
Oy 454 PARGLEUHIISATRGINLEUTHRTYASPLLEUHPRO----- 466
Db 89872 CCGGTATACATGGGACGCTGACCTATGATCTGTCTCCCAAGTAGGCGTGAAGGATCAGGT 89931
Oy 466 ----- 466
Db 89932 GGTGGGCTGGGCTCAGGACCGCGCTAGTGCACAACCTCACTGTCTGTGTCTCTT 89991
Oy 467 -Thr-AspSerArgIAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGlu- 485
Db 89992 CAGCCGACAGTCCACAGACCTCGCCGCAACTAGTCACTATGCTTCCGTGACGAGG 90051
Oy 485 ----- 485
Db 90052 TGGGGGGGTGGGCTCGGGTTTGGAAAGTGCAGTCTGCCAGCTGCCAGTACTACCGT 90111
Oy 486 ----- 497
Db 90112 GCTTCCCAACCCCTTCACACTGACGATGACCGGACAAAGCTAGAGCTTCTTGGAGACC 90171
Oy 498 ThrPheLeuLysTyrArgGlyThrGlnAla 507

```

```

Db 90172 ACTTTCTCAAGTACCTGGGACACAGCG 90201
RESULT 14
LOCUS BC074684
DEFINITION Xenopus tropicalis nuclear receptor binding protein, mRNA (cdna
ACCESSION BC074684
VERSION BC074684.1
KEYWORDS GI:49250535
SOURCE MGC.
ORGANISM Xenopus tropicalis (Silurana tropicalis)
Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 3046)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heist, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.S., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Miliani, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.U., Malek, U.A., Gunaratne, P.H., Richards, S.,
Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kietzman, M., Maman, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J., and Marz, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3046)
Klein, S. and Gerhardt, D.S.
Direct Submission
Submitted (23-JUN-2004) National Institutes of Health, Xenopus Gene
Collection (XGC) National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: gcapbs-remail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: DOE Joint Genome Institute, Walnut Creek,
California
Web site: www.jgi.doe.gov
FEATURES
Source
1..3046
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/cclone="MGC:69507 IMAGE:5336424"
/issue_type="Embryo, Silurana tropicalis, tailbud (stages
20-27)"
/cclone_id="NICHD_XGC_Emb7"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 130 Row: C Column: 1
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.
 Location/Qualifiers
 1..3046

	ORGANISM
	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE AUTHORS	1 (bases 1 to 2128) Straussberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Schmeier C.M., Sculler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K., Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F., Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheer T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raja S.S., Iqbalano N.A., Peters G.J., Abramson R.D., Millahy S.J., Bosak S.A., McSwan P.J., McKernan R.J., Malek J.A., Gunaratne P.H., Richards S., Woley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shcherchenko Y., Bonfield G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield J.S., Krzywicki M.I., Skalka U., Smalls D.E., Schmer A., Schein J.R., Jones S.J. and Marra M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
TITLE	12477932
JOURNAL	2 (bases 1 to 2128)
PUBMED	Straussberg R.
REFERENCE	Direct Submission
AUTHORS	Submitted (21-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
JOURNAL	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: gcgaps-remail.nih.gov Tissue Procurement: Lochar Hennighausen Ph.D., Chu-Xia Deng Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amgdbcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulik, S.W., Louisedge, H., Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanaevii, A.N., Gibbs, R.A.
REMARK	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Series: IRAK Plate: 6 Row: a Column: 1 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22219433. Location/Qualifiers
FEATURES	1..2128
SOURCE	/organism="Mus musculus" /mol_type="mRNA" /strain="129, C57BL/6J, FVB/N" /db_xref="taxon:10099" /clone="WGC:6961 IMAGE:3154089" /tissue_type="mammary tumor. Brca1-/fl; MMTV-Cre model. 10 months old, gross tissue." /clone_lib="NCI CGAP_Mam3" /lab_host="MDH10B" /note="vector: pcMV-SPORT6"
gene	1..2128 /gene="Nrbp" /note="synonym: MGC6961" /db_xref="GeneID:192292" /db_xref="MG1:2183436" 75..1682 /gene="Nrbp"
CDS	

	/codon_start=1	
	/product="nuclear receptor binding protein"	
	/protein_id="AAH04756.1"	
	/db_xref="GI:13435804"	
	/db_xref="GeneID:192292"	
	/db_xref="MGI:2183436"	
	/translation="MSRSGSQVSSGSDPKYESSLAPGLTSPSPVSTTSASR EESESEDESEILIESPCGRWQKRRREVORNPGLDASYLAMDTGECEVMVEOER EKRYKKLOEKEKVAFVDNLIQLEHLINLYAKFHKYADVKNKARVITEVMSSGSLK FLKYKKKHKTMEKAMRWCKITCISALSYHSICDPRIHGNTCTPIFYONGLIK GSVAEDLTNNHYKTCREOKNHFPAPEEGEVNTVADIYSFGCALMAVLATLC NGESVVYPOEALISSAIOLEDSLOEPFOKCSOPARPPTBELFHPLFPVPIS LLAHCTYGHCHMTPEMLAEETTKMDTSAYLAELPAGRGREPVOTLYQSAPYLELD FLVEDRNQIYPLTAFLPRPOOQOEETSVPVPSVKPTPEPAVERTRKVLMOQ IESVEGVGNHLLTLKLKEDKNRLHSCDLMENESIPLDAELIVOGFTSEADOISR SLBETLNKFPTFRSTINTATVTSS"	
ALIGNMENT		
Alignment Scores:		
Pred. No.:	7,38e-122	Length: 2128
Score:	1537.50	Matches: 307
Percent Similarity:	73.15%	Conservative: 58
Best Local Similarity:	61.52%	Mismatches: 105
Query Match:	57.58%	Indels: 29
Dbl:	9	Gaps: 6
US-10-618-941-99 (1-507) x BC004756 (1-2128)		
OY	17	GIUATGUAARGLUSeRGluAepGluSeRApiLeleUGluSeRPProCYgLIyArg 36
DB	204	GAGGAAGAAGAAAGTAGAAGATAGACTTGACATCTTGAAAGATCACCCCTGGGGCG 263
OY	37	TTPGlnlyvArGaRGluGlInValAnglInLyAsmMeCProGlyLeugInseRthrPhe 56
DB	264	TGGCAAGAAGCGCGAAGAGGTGATCAGCCTTAATGTACCAGGCATTGACAGTGATAC 323
OY	57	LeuAlaMetAerThrGluGlInLyValGluValTPrAnsgluLeuNHsPhRegLyAap 76
DB	324	CTGGCGCATGAGTACAGAGGAAGGTGTGGAGTTGTGTGAATGAGGATCAGTTCTGTGA 383
OY	77	ArgLyAlaPheAlaAlaNHsaglUGluLySILEgInThrValPheglUGluInleuValLeu 96
DB	384	CGCAAGAACTACAACCTGACAGGAGAAAGAGGCCGTCGACGTGTTATATATCTGATTCAG 443
OY	97	ValAspNHsPProAsnLLeValLyLeuNHsLySlyTYTrPLeuAapThrSerGluAlaCy9 116
DB	444	CTGGAACTCTTAATCACTGTTTAAGTTTACAAATATTTGGGCGTGATGTTAAAGAACAAAG 503
OY	117	AlaArgValLLePHeLLeThrgLUtyrValSerSerGlySerLeuLyagLinPheLeuLy9 136
DB	504	GCTAGGGGTGATTTTCATACACAGAAATAACATGTCCTCTGGGAGTCCTTAAGCAGTTTCTGAG 563
OY	137	LysThTyLyvLvAsNHsLyLeuAlaMerAenAlaArgAlaTPrLyvArgTrTryCyvTHnglm 156
DB	564	AAGACCAAAGAAACACACAGACTATGTAATGAAGAGGCTTGAAAGGCTGGGTGTACACAG 623
OY	157	IleLeuSerAlaLeuSerPheLeuNHsAlaCy9SerProPoiLLeIHsGLyAsmLeu 176
DB	624	ATTCCTCTGCGCTTAAGTACTGACATCTCTGTGACCTCCCATCATTCATGGGAACCTG 683
OY	177	ThrSerAapThrLLePHeLLeGlnNHsAnglLyLeuLLeLyLLeGlySerValTPrNHs 196
DB	684	ACCTGTGACACCATCTTCATCCAGACAAACGAGCATCAAAATGAGCTCTGTG----- 737
OY	197	ArgLLePHeSerAenAlaLeuArgProPThrAlaLeuProAspAerleuAygSerPro 216
DB	738	-----GCTCTGACACCTATCAACAATTCAC 761
OY	217	IleArgAlaGluArgGluGluLeuArgAsmLeuNHsPhePheProProGlyTyrgLyGlu 236
DB	762	GTGAAGACTTCCCGGGAAGAACAGAAACCTTCACTTTTTTGGACCAAGATTTGAGAA 821
OY	237	ValAlaAap---GlyTyThrAlaValaAspLLePHeSerPheGlyMetCyvalaleuGlyMet 255

```
Db      822  GTCCACAACAGTCAGACAGAGATCTACTCTTGGCATGTGTGCATGGAGATG 881
Qy      256  A1aValLeuGlu1IeGIInThraENG1uAerThArg---ValThrGluGluA1a1eA1a 274
Db      882  GCAGTGTGAGATTGAGGCAATGGCAATGGCTCATATATGTGCACAGAGAGCCATCAGC 941
Qy      275  ArgAlaArgHisSerLeuSerLeuSerProAmeMetArgGluPhe11eLeuGysCysLeuA1a 294
Db      942  AGTGCATATCCAGCTACTATGAAAGCTATTACAGAGGAGATTATTCAAAAGTGCCTGCAG 1001
Qy      295  ArgAerProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
Db      1002  TCTGAGCTGTGCTGGAGACCAACAGCAGAGAACTTCTGTCCACCCAGACGTGTGAA 1061
Qy      315  ValHisSerLeuLeuLeuA1a1aHisCysPhe11eGlnHisGlnThrLeuMetPro 334
Db      1062  GTGCCCTCACTCAAGCTTCTGTGCTCACTGTATGTGTGGGCAACACATGATCCCA 1121
Qy      335  GluAerValValGluGluThyValMetAerLeuHisValValLeuA1aGluLeu 354
Db      1122  GAGAGCTGTAGAGAGATACACAAAGACATGATACCAAGTGTCTACTAGCTGAAT 1181
Qy      355  ProArg---ProArgArgProProLeuGlnThrArgTyrSerGluValSerPheMetGlu 373
Db      1182  CCGCAGAGGCGCAGAGAGACAGATTCAGACTTGTACTCTCAGTACACAGCCCTAGA 1241
Qy      374  LeuAerPheLeuGluAerValArgAENG1yleTyrProLeuMetAerPheA1a1a 393
Db      1242  TTAGACAAATTCCTTGAAGATGCAGGAATGGGATCTACCTCTGACAGCCTT----- 1295
Qy      394  ThrArgProLeuGluLeuProArgValLeuA1aProProGluGluVal----- 410
Db      1296  -----GGCTACCTCGGCTCTCAGAGCCACAGCAGAGGAGGTGACATCACCT 1343
Qy      411  -----GlnValAlaLeuThrProThrProGluProPheAerSerGluThrArgGlys 427
Db      1344  GTTGGCCCCCTGTCTCAAGACTCCACTCTGAGCCAGCTGAAGTGAAGACAGAAAG 1403
Qy      428  Val11eGlnMetGlnCysAerLeuGluArgSerGluAerPheValArgTTrpHisLeuThr 447
Db      1404  GTGTGTCTGATGCAGTCAATCGAATCTGTGAGAGAGGAGTCAACACCATCTAACA 1463
Qy      448  LeuLeuLeuValLeuGluAerArgLeuHisArgGlnLeuThrTyrAerLeuLeuProThr 467
Db      1464  CTTCTGTGAAGCTGAGAGCAAAATTGAACCGGACCTGAGCTGATGCCCAAT 1523
Qy      468  AerpSerAlaGlnAerPheA1aSerGluLeuValHisTyrGlyPheLeuHisGluAerAer 487
Db      1524  GAGAGCATCCCGACTTGGCAGCTGAGCTGTGCACTGGGCTTCATTAGTAGGCTGAT 1583
Qy      488  ArgMetLysLeuA1a1aPheLeuGluSerThrPheLeuLysTyrArgGlyThrgln 506
Db      1584  CAGAGCGGCTGACTTCTGTGCTGAGAGAGCGCTCAACAAGTTCAACTTCAACAGG 1640
```

Search completed: January 13, 2006, 23:36:28
Job time : 6726 secs

THIS PAGE BLANK (USP10)

QY	1	MetAlaAlaProGluProAlaProAlaProAlaAlaArgGluAlaArgGluAlaArgGluAlaArgGluAla	20
Db	282	ATGGGGCCCCCGGAGCCCGGCGCGGAGGGGGCGCCGAGAAACGGGAAACGGGAAACGGGAAACGGGAGGAC	341
QY	21	GluSerGluAlaSerGluSerAspIleLeuGluGluSerProCysGlyArgTrpGlnIlyAspArg	40
Db	342	GAGAGCGAGAGACGAAAGGCAACTCTTGAGAGAAACCCCTGTGTCTCGCTGGAAAGCA	401
QY	41	ArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnInsertThrPheLeuAlaMetAsp	60
Db	402	CGGGGCGAGGTAAACCAAGGGAACAATGCCAGGGCTTCCAGAGCACTTCTCAGCCATGGAC	461
QY	61	ThrGluGluGluValGluGluValValTrpAsnGluLeuAsnIleAspGluValAspArgCysAlaPhe	80
Db	462	ACGGAGAGGGGGGTAGAGGTGTGTGGACAGAGCTCACTTCCAGACAGAAAGGCTTCC	521
QY	81	AlaAlaAsnIleGluGluIlyLeuIleGlnThrValPheGluGlnLeuValIleValAsnIlePro	100
Db	522	GCGGGCGCACAGAGAAAGATCCAGACCGGTTCGAGACAGCTGTGTCTGTGTGGACCAACCG	581
QY	101	AsnIleValIlyLeuAsnIleValIleThrIleAspTrpIleSerGluAlaCysAlaArgValIle	120
Db	582	AACATCGTAAGTTGCACAAAGTACGGCTGGATACCTCGAGGCTCGCGAGGGGTCAATC	641
QY	121	PheIleThrGluTrpValIleSerSerGlySerLeuIlyGlnPheLeuIlyAsnIleThrIlyAsnIly	140
Db	642	TTTCATCACAGAGTACGTGTCAATCAAGAGCGCTCAAGCAATCTCTAATAAAGCAAGAG	701
QY	141	AsnIleValIlyAlaMetAsnAlaArgAlaTrpIlyValArgTrpCysThrGlnIleLeuSerAla	160
Db	702	AACCCAAAGGCGCATAGAACGCCGGGCGCTGGAAAGCCCTGGTGACCAACATCTGTCTTCG	761
QY	161	LeuSerPheLeuAsnIleAlaCysSerProProIleIleAsnIleGluAsnLeuThrSerAspThr	180
Db	762	CTCAGCTTCTCTGACAGCTCGACGCCCCCAATCACTCAAGGGAACTTACACAGGACACC	821
QY	181	IlePheIleGlnIleAsnGlyLeuIleIlyValIleGlySerValTrpIleAspGluIlePheSer	200
Db	822	ATCTTCATTCAGACAAAGCGCTCAAGATCGGCTCGTGTGGACCAAGATCTTCTCC	881
QY	201	AsnAlaLeuAspProProThrAlaLeuProAspAspIleAspSerProIleArgAlaIleIu	220
Db	882	AAT-----GCATCTTCAGATGATCTCCAGAGGCCCATCCGCGCTAG	923
QY	221	ArgGluGluLeuAspArgAsnLeuAsnIleAspPheProProGluTrpGlyGluValAlaAspGly	240
Db	924	CGAGAGGAACCTTCGGAACCTGCACCTTCCCCCCAGAGTATGAGAGAGTGGCCGATGG	983
QY	241	ThrAlaValAspIleAspSerPheGlyMetCysAlaLeuGluMetAlaValIleGluIle	260
Db	984	ACCGCTGTGGACATCTTCTCTCTGGATGTGTGTGGCTGGAGATGGCTGTACTGGAAATC	1043
QY	261	GlnThrAsnGlyAspTrpArgValIleThrGluGluAlaIleAlaAspAlaArgIleSerIleu	280
Db	1044	CAGACCAATGGGGAGAACCCCGGATCAAGAGAGAGGCATTTCTCGGCCAGAGGACCTCGCTG	1103
QY	281	SerAspProAsnMetArgGluPheIleLeuCysValaLeuAlaArgAspProAlaArgArg	300
Db	1104	AGTGACCCCAACATCGGGAGAGTTCACTCTTTGTCTGCGCCGGGACCTTCGCCGCGG	1163
QY	301	ProSerAlaAsnIleSerLeuLeuPheIleArgValIleuPheGluValIleIleSerIleuIlyLeu	320
Db	1164	CCCTTGCCCAAGGCTCTCTCTTCCACGCGGCTCTTCAGAGTCACTCGGTGAAGCTC	1223
QY	321	LeuAlaAlaAsnIleCysPheIleGlnIleGlnIlyTrpLeuMetProGluAsnValValGluGlu	340
Db	1224	CTGGAGGCCACTGTCTTATCCAGACCAAGTACCTCAATGCTGAAATGTGTGTGAGGAG	1283
QY	341	IlyThrIlyValAlaMetAspLeuAsnIleAlaValLeuAlaGluLeuProArgTrpArgArgPro	360
Db	1284	AAGACCAAGGCGCATAGACTGTGACCGGCTTGTGGAGAGTTCCCGGCGCCGCGAGGCGCC	1343

Qy	361	ProLeuGIINTPrAqTYSerGIuValSerPhneMeGIuLeuAspYsPhneGIuAsp	360
Db	1344	CGGTGCAGTGGCGGGACTCGGAAGCTCTTCCTCANGAGCTGGACAAATTCCTGGAGGAT	1404
Qy	381	ValATgaSeGIYIleTYrProLeuMeAenPhneAaIaThrArpProLeuGIYLeuPro	400
Db	1404	GTCAAGAAATGAAATCTACCACTGATGAACCTTTGGAGCCACTGACCCCTGGGGCTGCC	1464
Qy	401	ArgValLeuAaIaProProBGIuGIuValGIuLYsAaIaLYsThr-ProThrProGIuPr	420
Db	1464	CGTGTGCTGGCCCCACCCCCGGAGAGGTCCAAAAGCCAAAGACCCCCGACAGAGCC	1524
Qy	420	oPhaSePeSeRGITuThrArgLYsValIleGIuMeGIuCYsAenLeuGIuArSeGIuAs	440
Db	1524	CTTTACTCTTGAGACCAAGAAAGTCATCCAGATGCAAGTGCACCTGAGAGAAAGCCGAGGA	1584
Qy	440	PLysAaIaArgTTPHISLeuThrLeuLeuValLeuGIuAspArgLYsAaIaArgLYuLe	460
Db	1584	CAGGCGCGCTGGCATCTCACTCTGCTTGGTGTCTGGAAAGACCGGCTGCACCGGAGCT	1644
Qy	460	uThrTYrAspLeuLeuProThrAspSeRaIaGIuAspLeuAaIaSeGIuLeuValHisTY	480
Db	1644	GACCTACGACCTGTCTCCCAACGAGACAGCGCCCAAGACCTCGCTCGAGCTCGGACCTA	1704
Qy	480	rGIYsPhneLeuHISGIuAspAspArgLYsValAaIaPhneLeuGIuSerThrPhneLe	500
Db	1704	TGGCTTCTCTCCACCGAGAGACCGGATGAAAGCTGGCGGCTTCTCTGGAGACACTTCCCT	1764
Qy	500	uLYsTYrArgGIYThrGIuAaIa 507	
Db	1764	CAAGTACCTGGGAGACCCAGGCC 1785	
	RESULT 2		
	US-09-345-473E-9		
	/ Sequence 9, Application US/09345473E		
	/ Patent No. 6558903		
	/ GENERAL INFORMATION:		
	/ APPLICANT: Hodge, Martin		
	/ TITLE OF INVENTION: No. 6558903el Kinases and Uses thereof		
	/ FILE REFERENCE: 35800/183781		
	/ CURRENT APPLICATION NUMBER: US/09/345,473E		
	/ CURRENT FILING DATE: 1999-06-30		
	/ NUMBER OF SEQ ID NOS: 62		
	/ SOFTWARE: FastSeq for Windows Version 4.0		
	/ SEQ ID NO 9		
	/ LENGTH: 981		
	/ TYPE: DNA		
	/ ORGANISM: Homo sapiens		
	/ FEATURE:		
	/ NAME/KEY: misc_feature		
	/ LOCATION: (1)...(981)		
	/ OTHER INFORMATION: n = A,T,C or G		
	US-09-345-473E-9		
	Alignment Scores:		
	Pred. No.:	2,52e-169	981
	Score:	1617.00	Matches: 312
	Percent Similarity:	96.89%	Conservative: 0
	Best Local Similarity:	96.89%	Mismatches: 3
	Query Match:	60.56%	Indels: 7
	DB:	3	Gaps: 1
	US-10-618-941-99 (1-507) x US-09-345-473E-9 (1-981)		
Qy	1	MeTAaIaProGIuProIaProArGaRgAaIaArgGIuArGIuArGIuArGIuAsp	20
Db	35	ATGGGGGCCCCCGAAGCGGGCGCGAGCGCGGCCGGGAACGGGAGCGGAGCGGAGAC	94
Qy	21	GIuSeRGiUaSPGIuSeRaSPiIleuGIuGIuSeProCYseGIYArGTpGIuLYsArg	40
Db	95	GAGAGCGAGAGACGAGACGACACTCTCTGGAGGAAAGCCCTGTGTGTGTGTGGCAAAAGCA	154
Qy	41	ArgGIuGIuValaSnGIuCYsAenMeTProGIYLeuGIuInSerThrPhneAaIaMeTAsp	60

```

Db      155 CGGAGCAGGTAAACCAAGGAAACATCCAGGGCTTCAAGACACCTTCTAGCCATGAGAC 214
Qy      61  ThrGluGluGlyValGluValValTTPaangluleuHsPheGlyAspArglyAspAlaPhe 80
Db      215 ACCGAGAGAGGGGCTGAGAGGTGGTGTGAAAGAGCTTCACTCGGAGACAGAAAGGCTTTC 274
Qy      81  AlaAlaHsIsgluGluValIleGlnThrValPheGluGlnLeuValLeuValAspHisPro 100
Db      275 GCGGGGCAACGAGAGAAAGATCCAGAACCGGTTCGAGCACTGGTGTGGTGAACCAACCGG 334
Qy      101 AsnIleValIleLeuHsIleYsYrTTPleuAspThrSerGluAlaCysAlaArgValIle 120
Db      335 AACATCGTGAAGTTCGACAAATACCTGGCTGATACCTCGAGGCTCGCGAGGCTCATC 394
Qy      121 PheIleThrGluTyrValSerSerGlySerIleuValGlnPheLeuValYsYrThrIleYs 140
Db      395 TTCATCACAGATAGCTGTCTATCAGGACCTCAAGCAATTCCTCAAAAAGCAAGAGAG 454
Qy      141 AsnHsIleValIleMetAsnAlaArgAlaTATpIyAspArgTTPCysThrGlnIleLeuSerAla 160
Db      455 AACCAACAGGCAATGAACCCCGGCTGAAAGCGCTGGTGCACGCAATCCTGTCTGCG 514
Qy      161 LeuSerPheLeuHsAlaCysSerProProIleIleHsGlyAsnLeuThrSerAspThr 180
Db      515 CTCACCTTCTGCAAGCTGCAAGCCCGCAATCATCAGGAACTGACCAAGCAACAC 574
Qy      181 IlePheIleGlnHsIleAsnGlyLeuIleYsIleGlySerValTTPHisArgIlePheSer 200
Db      575 ATCTTCATTCAGCAACAACGCGCTCATCAAGATCGGCTCGGTGGCAACGAATCTTCTCC 634
Qy      201 AsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArgIleGlu 220
Db      635 AAT-----GCACTTCCAGATGATCTCGAAGCCCAATCCGCGCTGAG 676
Qy      221 ArgGluGluLeuArgAsnLeuHsIlePhePheProProGluTyrGlyGluValAlaAspGly 240
Db      677 CGAGAGGAACCTCGGAACCTGCACTTCTCCCGCAAGATAGAGAGGTGGCCCATGGG 736
Qy      241 ThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeuGluIle 260
Db      737 ACCGCTGTGACATCTTCTTCTTGGAGATGTGCGCTGAGATGGCTGTACTGGAATTC 796
Qy      261 GlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeu 280
Db      797 CAGACCAATGGGAAACCCGGGTCAACAGAGAGGCCATTGCTGGCCAGGCACTCGCTG 856
Qy      281 SerAspProAspMetArgGluPheIleLeuCysCysLeuAlaArgAspProAlaArgArg 300
Db      857 AGTGACCCCAACATCGGAGATTCATCTTGTGCTGTGGCCGGGACCTGCGCGACG 916
Qy      301 ProSerAlaHsIleSerLeuLeuPheHis-ArgValLeuPheGluValHsIleSerLeuYsIle 320
Db      917 CCCTCTGTGCAACAGCTCTCTTCCACNCGGCTGCTTNGAGGTGCACTGCTGAAGCT 976
Qy      320 uleu 321
Db      977 CCTG 980

```

```

RESULT 3
US-09-862-027-9
Sequence 9, Application US/09862027
Patent No. 6858418
GENERAL INFORMATION:
APPLICANT: Hodge, Martin R.
TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof
FILE REFERENCE: 35800/234862
CURRENT APPLICATION NUMBER: US/09/862,027
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/345,473
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 82
SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 9
; LENGTH: 981
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (981)
; OTHER INFORMATION: n=A,T,C or G
US-09-862-027-9

Alignment Scores:
Pred. No.: 2,52e-169 Length: 981
Score: 1617.00 Matches: 312
Percent Similarity: 96.89% Conservative: 0
Best Local Similarity: 96.89% Mismatches: 3
Query Match: 60.56% Indels: 7
DB: 3 Gaps: 1

US-10-618-941-99 (1-507) x US-09-862-027-9 (1-981)
Qy      1  MetAlaIaAProGluProAlaProArgAlaArgGluArgGluArgGluArgGluArg 20
Db      35  ATGGGCGGCCCGGAGCCCGCGCGAGCGCGCGCGGAAACGGAGCGGAGCGGAGAGC 94
Qy      21  GluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyValTTPGlnYsArg 40
Db      95  GAGACGAGAGAGAGAGAGCAATCTTGAGAGAAAGCCGTGTGTGCTGGCAAAAGCGA 154
Qy      41  ArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
Db      155  CGGAGCAGGTAAACCAAGGAAACATGCCAGGGCTTCAGAGCACTTCTTACCATGGAC 214
Qy      61  ThrGluGluGlyValGluValValTTPaangluleuHsPheGlyAspArglyAspAlaPhe 80
Db      215  ACCGAGAGAGGGGCTGAGAGGTGGTGTGAAAGAGCTTCACTCGGAGACAGAAAGGCTTTC 274
Qy      81  AlaAlaHsIsgluGluValIleGlnThrValPheGluGlnLeuValLeuValAspHisPro 100
Db      275  GCGGGGCAACGAGAGAAAGATCCAGAACCGGTTCGAGCACTGGTGTGGTGAACCAACCGG 334
Qy      101 AsnIleValIleLeuHsIleYsYrTTPleuAspThrSerGluAlaCysAlaArgValIle 120
Db      335  AACATCGTGAAGTTCGACAAATACCTGGCTGATACCTCGAGGCTCGCGAGGCTCATC 394
Qy      121 PheIleThrGluTyrValSerSerGlySerIleuValGlnPheLeuValYsYrThrIleYs 140
Db      395  TTCATCACAGATAGCTGTCTATCAGGACCTCAAGCAATTCCTCAAAAAGCAAGAGAG 454
Qy      141 AsnHsIleValIleMetAsnAlaArgAlaTATpIyAspArgTTPCysThrGlnIleLeuSerAla 160
Db      455  AACCAACAGGCAATGAACCCCGGCTGAAAGCGCTGGTGCACCAAGATCCTGTGCTGG 514
Qy      161 LeuSerPheLeuHsAlaCysSerProProIleIleHsGlyAsnLeuThrSerAspThr 180
Db      515  CTCACCTTCTGCAAGCTGCAAGCCCGCAATCATCAGGAACTGACCAAGCAACAC 574
Qy      181 IlePheIleGlnHsIleAsnGlyLeuIleYsIleGlySerValTTPHisArgIlePheSer 200
Db      575  ATCTTCATTCAGCAACAACGCGCTCATCAAGATCGGCTCGGTGGCAACGAATCTTCTCC 634
Qy      201 AsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArgIleGlu 220
Db      635  AAT-----GCACTTCCAGATGATCTCGAAGCCCAATCCGCGCTGAG 676
Qy      221 ArgGluGluLeuArgAsnLeuHsIlePhePheProProGluTyrGlyGluValAlaAspGly 240
Db      677  CGAGAGGAACCTCGGAACCTGCACTTCTCCCGCAAGATAGAGAGGTGGCCCATGGG 736
Qy      241 ThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeuGluIle 260
Db      737  ACCGCTGTGACATCTTCTTCTTGGAGATGTGCGCTGAGATGGCTGTACTGGAATTC 796
Qy      261 GlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeu 280

```

```

Db      797  CAAACCAATGGGGACACCGGGTCCACAGAGAGGACCATTCCTCC
QY      281  SeASpProxasmMeArGluPheIleLeuCySeYseu1aa
Db      857  AGTGAACCCCAACATCGGGAGTTCATCTTTCGCTGCGCCGCC
QY      301  ProSeRnIahisSerLeuLeuPheHis-ArGValLeuPheGluV
Db      917  CCCTGTGTCCACAGCCTCTCTTCCACGCGGTCTCTTGAGAG
QY      320  uLeu 321
Db      977  CCTG 980

RESULT 4
US-08-933-750C-69
; Sequence 69, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; City: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/933,750C
; FILING DATE: September 23, 1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0356 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 69:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2029 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAITUT03
; CLONE: 864683
; US-08-933-750C-69

Alignment Scores:
Pred. No.:      1,13e-159      Length:      2029
Score:          1534.50       Matches:     303
Percent Similarity: 73.54%    Conservative: 59
Best Local Similarity: 61.62% Mismatches:     102

```

Query Match:	57.47%	Indels:	29
DB:	2	Gaps:	6
US-10-618-941-99 (1-507) x US-08-933-750C-69 (1-2029)			
QY	17	GLUARGGLUABRGGLUSeRGLUABRGGLUSeRkPpLleuGLUGLUSeRProCYeGLYARG	36
DB	240	GAGGAGAAAGAAAGAAAGTGAAGTGAAGTCTGATTTTGGAAAGAGTCCCTGTGGGCGC	299
QY	37	TRPGLUYSARGATGSLUGLValAsnGLInGLYAsnMetProGYLeuGLInSeRThrPhe	56
DB	300	TGGCAGAGAGCCGAGAGAGGTGAATTAACGGAAATGACAGGATTAAGACAGTGCATAC	359
QY	57	LeuAlaMetAspThrGLUGLUGLValGluValValTPraenGLInLeuHsPheGLYAsp	76
DB	360	CTGGCCATGATACAGAGAGAGGTGAAGTGTGTGTGAATTAAGATGATCAGTTCCTGAA	419
QY	77	ArgGLYAlaPheAlaAlaHsGLUGLUYSLeGLInThrValPheGLUGLInValLeu	96
DB	420	CGCAAGAACTCAAGCTGCAGAGAGAAAGGTGTGCTGTGTGTGAATTAACCTGATCAA	479
QY	97	ValAspHsProAlaValLeuLeuHsAlaCysSerProPoiLeHsGLYAsnLeu	116
DB	480	TTGGAGCATCTTAACATTTGTAAGTTTCAAAATTAATGGCTGCAATTAAGAGAACAG	539
QY	117	AlaArgValLlePheLleThrGLUTyrValSerSerGLYSerLeuYSGLInPheLeuYS	136
DB	540	GCCAGGCTCATTTTATACAGAAATACATGTCATCTGGAGTCTGAGCAATTTCTGAG	599
QY	137	LYSThrLYSLYValAsnHsLYSLAlaMetCAsnAlaArgAlaTPLYValGTPCYSThrGLIn	156
DB	600	AAGACCAAAAGAAACCCACMACCATGATTAAGAAAGGATGAAAGCTGTGTGCACACAA	659
QY	157	IleLeuSerAlaLeuSerPheLeuHsAlaCysSerProPoiLeHsGLYAsnLeu	176
DB	660	ATCTCTTGCCCTTAAGCTTACCTGACTCCTGTGACCCCTCATTCATCAAGGAACTG	719
QY	177	ThSerAspTrLlePheLleGLInHsAsnGLYLeuLleYSLeGLYSerValTPHs	196
DB	720	ACCTGTGACACATCTTTCATCACAACAAGCATCANTCAAGATTGCTCTGTG----	773
QY	197	ArgLlePheSerAsnAlaLeuArgProProThrAlaLeuProAspPheLeuArgSerPro	216
DB	774	-----GCTCTGTGACATCAATCAACATCAT	797
QY	217	IleArgAlaGLUARGGLUGLInLeuArgAsnLeuHsPhePheProProGLInTYGLIn	236
DB	798	GTGAAAGACTTGTGAGAAAGACAGAAAGATTTAACACTTTTGACCAAGATATGAGAA	857
QY	237	ValAlaAsp---GLYThrAlaValAspLlePheSerPheGLYMetCysAlaLeuGLInMet	255
DB	858	GTCACATAATGTGACAAACAGCAGTGCATCTACTCTTTGGCATGTGTGCACCTGAGATG	917
QY	256	AlaValLeuGLInLleGLInThrAsnGLYAspThrArg---ValThrGLInAlaLleAla	274
DB	918	GCAGTGTGAGATTCAGGCGAAATGGAAGTCTCATATGTGCCACAGGAAGCATTCAGC	977
QY	275	ArgAlaArgHsSerLeuSerAspProAsnMetArgGLInPheLleLeuCysCyleuAla	294
DB	978	AGTGCATCCAGCTTCTTAAGAACCCATTACAGAGGGAGTTCATTAAGATGCTTGAG	1037
QY	295	ArgAspProAlaArgAspProSerAlaHsSerLeuLeuPheHsArgValLeuPheGLIn	314
DB	1038	TCTGAGCTGCTCGCAGACCAACAGCCAGAGAACTTCTGTCCACCCAGCATTTGTTGAA	1099
QY	315	ValHsSerLeuLYSLLeuLeuAlaAlaHsCysPheLleGLInHsGLInTYrLeuMetPro	334
DB	1098	GTCGCCCTGTGCAAACTCTTGCGGCCCACTGCAATTTGGGAGACAAACATGATCCCA	1157
QY	335	GLUAsnValValGLUGLUYSThrLYSLAlaMetCAspLeuHsAlaValLeuAlaGLInLeu	354
DB	1158	GAGAAAGCTCTTAAGAGAGATCACCAAAACATGATACTAGTGCCTGATCTGGTGAATCT	1217

Oy	355	ProArg-----ProArgArgProProleuGlnIntPrpaGlyTyrSerGluValSerHeMetGlu	373
Db	1218	CCTGAGAGACCAGAGAGAGAACCACTTGACACTTGTACTCTTCAGACACCACTCTGGAA	127
Oy	374	LeuApblyPheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnDheAla	393
Db	1278	TTAGATAAATCTCTGAAGATGTCAAGATGGAGATCTATCTCTGCACAGCTTT-----	133
Oy	394	ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal-----	410
Db	1332	-----GGGCTGCTCTGGCCCCAGCAGCAGCAGAGAGAGGTGCATCACCCT	137
Oy	411	-----GlnValAlaTyrThrProThrProGluProPheAspSerGluThrArgLys	427
Db	1380	GTCGTGCCCCCTCTGTCAAGACTCGACACCTGAAACAGCTGAGGTGAGACTCGCAAG	143
Oy	428	ValIleGlnMetGlnCysAsnLeuGluLutArgSerGluAspLysAlaArgTrpHisLeuThr	447
Db	1440	GTGGCTGCAATGCACTGCACATTGACTCGGTGGAGAGGGAGTCAACACCACTGCACA	149
Oy	448	LeuLeuLeuValLeuGluAspArgPheHisArgGlnLeuThrTyrAspLeuLeuProThr	467
Db	1500	CTTCCTGCTGAAGTTGAAGACAACTGAAACCGGCACTTGAGCTGTGACTGATGCCAAAT	155
Oy	468	AspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAsp	487
Db	1560	GAGAAATATCCCGGAGATGGCGGCTGAGCTGGTGCAGACTGGGCTTCATTAGTAGGCTGAC	161
Oy	488	ArgMetLysLeuAlaAlaPheLeuLutSerThrPheLeuLysTyr	502
Db	1620	CAGAACCCGATGACTTCTCTGCTTAAGAGACCTTGAAACAATTC	1664

US-09-234-613-69
Sequence 69, Application US/09234613
Patent No. 6132973
GENERAL INFORMATION:
APPLICANT: Lal, Preeti
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Shah, Purni
APPLICANT: Au-Young, Janice
APPLICANT: Yue, Henry
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
NUMBER OF SEQUENCES: 98
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/234,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/933,750
FILING DATE: September 23, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166

```

?       TELEX:
?       INFORMATION FOR SEQ ID NO: 69
?       SEQUENCE CHARACTERISTICS:
?       LENGTH: 2029 base pairs
?       TYPE: nucleic acid
?       STRANDEDNESS: single
?       TOPOLOGY: linear
?       IMMEDIATE SOURCE:
?       LIBRARY: BRALTTU03
?       CLONE: 864683
?       US-09-234-613-66

```

Alignment Scores:	
Pred. No.:	1.13e-159
Score:	134.50
Percent Similarity:	73.54%
Best Local Similarity:	61.62%
Query Match:	57.47%
DB:	3
Length:	2022
Matches:	305
Conservative:	59
Mismatches:	102
Indels:	29
Gaps:	6

US-10-618-941-99 (1-507) x US-09-234-613-69 (1-2029)

Oy	17	GIuIrvggIlaaPpGluSeGluaspGluSerApIleleuGIuGluSerProCysGIYArg	36
Db	240	GAGGAAGAAGAGAAAGTGAAGATGATGATCTGAGATTTTGAAAGAGTCCGCTGTGGGCC	239
Oy	37	TrpGluYsaTgATgAGluGluValaAsngInglYaAmMeCProGlyLeuGInSerThrPhe	56
Db	300	TGGCAGAAAGAGCCGAAAGAGGTGATCAACGAATGTATCCAGGATTTGACGTGATAC	359
Oy	57	LeuIaMeCAspThrGIuGluGlyValGIuValIValTPAsngIuLeuHiSpheGIaAsp	76
Db	360	CTGGCCATGAGATACAGAGAAAGTGTAGAGTGTGTGTGAATGAGCTACAGTTCTCTGAA	419
Oy	77	ArgIYsaIaPheAlaAlaHiSgluGluYsaIleGIInThrValPheGIuGluIneuValLeu	96
Db	420	CGCAAGAACTACAGCTCGAGGAGGAAAGGTTGGCTGTGTGTATGATATCATGATCAA	479
Oy	97	ValAspHiSPRoAsnIleValIYsaLeuHiSlySTYTrpLeuAspThrSerGIuaIaCys	116
Db	480	TTGAGAGCTCTTAACATGTGTTAAGTTTCAAAATTTGGCGTACATTTAAAGAACAAAG	539
Oy	117	AlaArgValIlePheIleThrGIuTYValSerSerGIYSerLeuIYsgInPheLeuYs	136
Db	540	GCCAGGGTCATTTTATACAGAAATACATGTCATCTGGAGTCTGAAGCAATTTCTGAAG	599
Oy	137	LYsThrIYsLYsaSniHLYsaIaMeCAsnIaAaGAlaTrpLYsaArgTrpCYsThrGIu	156
Db	600	AAGACCCAAAAAGAACCCACAAAGCGATGAATGAAGAAAGGCGATGGTGTCCACAA	659
Oy	157	IleLeuSerAlaLeuSerPheLeuHiSaIaCYsSerProGIIeIleHISglYaAsnLeu	176
Db	660	ATCCTCTTGCCCTTAAGCTACCTGGACCTCTGTGACCCCCCATCATTCATGGGAACCTG	719
Oy	177	ThiSerAspThrIlePheIleGIuHiSaSngIYLeuIleLYaIleGIYSerValTPHiS	196
Db	720	ACGTGTGACACCATCTTCATCCAGCAACGGAATCATCAAAATTTGGCTGTGG-----	773
Oy	197	ArgIlePheSerAsnIaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro	216
Db	774	-----GCTCCCTGACCTATCAACAATTCAT	797
Oy	217	IleArgAlaGIuArgGluGIuLeuArgAsnLeuHiSphePheProProGIuTYArgIu	236
Db	798	GTTAAAGACTTGTCGAGAAAGACAGAAAGATCTACACTTTTGACCCAAAGATGTGAGAA	857
Oy	237	ValAlaAsp---GIYThrAlaValAspIlePheSerPheGIYMeCYsaIaLeuGIuMeC	255
Db	858	GTCACATAATGTGACAAACAGCAGGTGACATCTACTCTTTGGCATGTGTCACTCGAGATG	917
Oy	256	AlaValLeuGIuIleGIInThrAsnGIYAspThrArg---ValIleGIuGIuAlaIleAla	274
Db	918	GCAGGTCTGAGATTCACAGGCAATGGAAGCTCTATATGTGCCACAGGAAGCATCTACAG	977

```
Oy 275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCyCysLeuAla 294
Db 978 AGTGCATCCAGCCTTCTAGAAAGACCATTTACAGAGGAGTTTCATCAAAAGGCTGGAG 1037
Oy 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
Db 1038 TCTGAGCCTGGCTGGAGACCAACAGCAGAGAACTTCTGTTCCACCCAGCATGTTTGA 1097
Oy 315 ValHisSerLeuValLeuLeuAlaAlaHisCysPheIleIleGlnHisGlnTyrLeuMetPro 334
Db 1098 GTGCCCTGCTCAAACTCTGGCGGCCCATCTGCATCTGGGACACCAACATGATCTCCA 1157
Oy 335 GluAsnValValGluGluValThrTyrValIleMetAspLeuHisAlaValLeuAlaGluLeu 354
Db 1158 GAGAAAGCTCTAGAGGAGATCACCAAAAACATGATCTAGTACCTGCGTCTGAAATTC 1217
Oy 355 ProArg---ProArgArgProProLeuGlnIleTyrArgTyrSerGluValSerPheMetGlu 373
Db 1218 CCTGCAGGACCAAGAGAGAACCACTTACACTTGTACTCTCAGTACCAAGCTCTGAA 1277
Oy 374 LeuAspLysPheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 1278 TTAGATAAATTCCTTGAAGATGTCAGAAATGGGATCTATCTCTGACAGCCTTT----- 1331
Oy 394 ThrArgProLeuGluValProArgValLeuAlaProProProGluVal----- 410
Db 1332 -----GGGCTCTCTGGCGGCCCAAGACACCAAGAGAGAGTACATCACCT 1379
Oy 411 -----GlnLysAlaLysThrProThrProGluProPheAspSerGluThrArgLys 427
Db 1380 GTCCGAGCCCTCTGTCAAGATCTCCGACCTGAAACCAAGCTGAGTGAAGATCTGGCAG 1439
Oy 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLysAlaArgTyrHisLeuThr 447
Db 1440 GTGGGCTGATGTCAGTGCACAACTTGAATGCGGTGGAGGAGGATCAAAACACACCTGACA 1499
Oy 448 LeuLeuLeuValLeuGluAspArgPheHisIleArgGlnLeuThrTyrAspLeuLeuProThr 467
Db 1500 CTCTGCTGAAGTTTGAGAGACAACTGAAACCGGACCTGAGCTGTGACTGATGCAAT 1559
Oy 468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAspAsp 487
Db 1560 GAGAAATATCCCGAGATTGGCGGCTAGCTGGTGCAGCTGGGCTTCATTAGTGAAGGCTGAC 1619
Oy 488 ArgPheLysLeuAlaAlaPheLeuGluSerThrPheLeuLysTyr 502
Db 1620 CAGAGCCGCTTGACTTCTCTGCTAGAGAGACCTTGAACAAGTTC 1664

RESULT 6
US-09-949-016-1428
; Sequence 1428, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1428
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1428
```

```
Alignment Scores:
Pred. No.: 1,26e-159 Length: 2163
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: 3 Gaps: 6

US-10-618-941-99 (1-507) x US-09-949-016-1428 (1-2163)
Oy 17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArg 36
Db 242 GAGAAAGAAAGAAAGAAAGTGAAGATGATCTGCAAGTTTGGAAAGACTCCCTCTGGCGGC 301
Oy 37 TrrGlnLysArgArgGluGluValAsnGlnLysAsnMetProGlyLeuGlnSerThrPhe 56
Db 302 TGGCAGAAAGAGGCGAAGAGAGTGAATCAACGGAAATGTACACAGGTAATTGACAGTGCATAC 361
Oy 57 LeuAlaMetAspThrGluGluGluValGluValValTrrAsnGluLeuHisPheGlyAsp 76
Db 362 CTGGCCATGTGATACAGAGAAAGGTGTAGAGGTGTGTGGAATGTAGATGACATTTCTGAA 421
Oy 77 ArgLysAlaPheAlaAlaHisGluGluLysIleGlnThrValPheGluGlnLeuValLeu 96
Db 422 CGCAAGAACTACAAAGCTGCAGAGAGAAAGGTGCTGCTGTGTTGATATCTGATTCAA 481
Oy 97 ValAspHisProAsnIleValLysLeuHisLysTyrTrrLeuAspThrSerGluAlaCys 116
Db 482 TTGAGCATCTTAACATTTGTTAAGTTTACAAATATTTGGGTGATTAAGAAACAAAG 541
Oy 117 AlaArgValIlePheIleThrGluTyrValSerSerGlySerLeuLysGlnPheLeuLys 136
Db 542 GCCAGGGTCATTTTATACAGAAATACATGTCATCTGGAGTGTGAAGCAATTTCTGAAG 601
Oy 137 LysThrLysLysAsnHisLysValIleMetAsnAlaArgAlaTrrLysArgTrrCysThrGln 156
Db 602 AAGACAAAAAGAAACCAAGACAGATGAATGAAGAAAGCAATGAGACCTGTGTCACAA 661
Oy 157 IleLeuSerAlaLeuSerPheLeuHisAlaCysSerProProlIleHisGlyAsnLeu 176
Db 662 ATCTCTCTGCCCCCTAACCTTACCTGCTGATGCCCTGCATTCATTCAGGAAACCTG 721
Oy 177 ThrSerAspThrIlePheIleGlnHisAsnGlyLeuLeuLysIleGlySerValTrrHis 196
Db 722 ACCTGTACACCACTTTCATTCACACAGAGACTCATCAAGATGCTGTGCTG----- 775
Oy 197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
Db 776 -----GCTCTGACATCATCAACATCAT 799
Oy 217 IleArgAlaGluArgGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGlu 236
Db 800 GTGAAGACTTGTCGAGAGAGAGCAAGAAATCTTACACTTCTTGGACACAGAGATGAGAA 859
Oy 237 ValAlaAsp---GlyTrrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMet 255
Db 860 GTCACTAATGTGACAAAGAGAGTGAATCATCTCTTGGCAGATGTGCACTGAGATGG 919
Oy 256 AlaValLeuGluIleGlnIleThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla 274
Db 920 GCACTGCTGAGATTCAGGGCAATGAGAGTCCATATATGTGCCACAGAAAGCATTCAGC 979
Oy 275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCyCysLeuAla 294
Db 980 AGTCCATCCAGCTTCTAAGAAAGCACTTACAGAGGAGATTCACTTCAAAAAGTGCCTGAG 1039
Oy 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
Db 1040 TCTGAGCCTGCTGGACCAACCAAGCAGAGAACTTCTGTTCCACCCAGCATGTTTGA 1099
Oy 315 ValHisSerLeuValLeuLeuAlaAlaHisCysPheIleIleGlnHisGlnTyrLeuMetPro 334
```

Db 1100 GTGCCCTGCTGAACTCCTTGGCGCCACCTGATTTGGGACACCAACATGATCCCA 1159
Qy 335 GUAenValaIGluuThyThyValaMetAspLeuH1AlaValaLeu1aGluLeu 354
Db 1160 GAGAAAGCTCTAGAGAGATCAACCAAAACATGATATCTAGTCCGCTACTGGCTGAAATC 1219
Qy 355 ProArg---ProArgArgProProLeuGlnTyrArgTyrSerGluValaSerPheMetGlu 373
Db 1220 CCTGACAGACCAAGAAAGAACACAGTTCAGACTTTTACTCTCAGTACACAGCTCGGAA 1279
Qy 374 LeuAspLysPheLeuGluAspValaArgaengly1IeTyrProLeuMetAsnPheAla1a 393
Db 1280 TTAGATTAATTCCTTGAAATGTCAGGAATGGATCTATCTCTGACACCTTT----- 1333
Qy 394 ThrArgProLeuGlu1LeuProArgValaLeuAlaProProProGluGluVala----- 410
Db 1334 -----GGGCTGCTGGCGCCCGCCAGCCACACAGAGAGAGGATGACATCCT 1381
Qy 411 -----GlnLysAlaLysThrProThrProGluProPheAspSerGluThraArgLys 427
Db 1382 GTGCTGCCCCCTCTGTCAAGACTCGACACCTGAACAGCTGAGGTGAGACTCGCAAG 1441
Qy 428 Val1IeGlnMetGlnCysaenLeuGluArgSerGluAspLysAlaArgTyrPhe1LeuThr 447
Db 1442 GTGCTGCTGATCAGTCAACATGATGCTGCTGAGAGAGAGATCAACACCACTGACA 1501
Qy 448 LeuLeuLeuValaLeuGluAspArgLeuH1aArgGlnLeuThraAspLeuProThr 467
Db 1502 CTTCGCTGAGATGAGAGCAACCTGAACCGGACCTGAGCTGACTGATGCCAAAT 1561
Qy 468 AspSerAlaGlnAspLeuAlaSerGluLeuValaH1aTyrGlyPheLeuH1aGluAspAsp 487
Db 1562 GAGATATCCCGGAGCTGGCGGCTGAGCTGCTGACCTGGCTTCTTATGTAAGCTGAC 1621
Qy 488 ArgMetLysLeuAlaAlaPheLeuGluSerThrPheLeuLysTyr 502
Db 1622 CAGAGCGGTTGACTTCTCTGCTAGAAAGACCTTGAAACAAGTTC 1666

RESULT 7
US-09-023-942A-9
Sequence 9, Application US/09023942A
Patent No. 6479274
GENERAL INFORMATION:
APPLICANT: (US only) ANTALIS Toni Marie and HOOPER John David
TITLE OF INVENTION: NOVEL MOLECULES
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 GARDEN CITY PLAZA
CITY: GARDEN CITY
STATE: NEW YORK
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/023,942A
FILING DATE: 13-FEB-1998
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: POS101/97
FILING DATE: 13-FEB-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP0422/97
FILING DATE: 18-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: International PCT Application
FILING DATE: 13-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: DIGIGLIO, FRANK S

REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 11168
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742 4343
TELEFAX: (516) 742 4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2241 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 166..1773
US-09-023-942A-9

Alignment Scores:
Pred. No.: 1,01e-156 Length: 2241
Score: 1508.50 Matches: 301
Percent Similarity: 72.93% Conservative: 60
Best Local Similarity: 60.81% Mismatches: 105
Query Match: 56.50% Indels: 29
Gaps: 6

US-10-618-941-99 (1-507) x US-09-023-942A-9 (1-2241)

Qy 17 GluArgLysAspGluSerGluAspGluSerAsp1IeLeuGluGluSerProCysGlyArg 36
Db 295 GAGGAAGAAGAAAGAAAGTGAATGATGCTGAGATTTTGGAAAGATCGCCCTGTGGGCGCC 354
Qy 37 TrrGlnLysArgArgGluGlnValaLenglnGlyAenMetProGlyLeuGlnSerThrPhe 56
Db 355 TGGCAGAAAGCGCGAAGAGGTGAATCAACCGAAATGATCCAGTATACAGTGCATC 414
Qy 57 LeuAlaMetAspThrGluGluGluValaGluVala1TrrPheGlnLeuH1aPheGlyAsp 76
Db 415 CTGGCCATGATACAGAGAGAGTGAAGGTTGAGTGTGGAATGAGATACATTCCTGTA 474
Qy 77 ArgLysAlaPheAlaAlaH1aGluGluLys1IeGlnThraPheGluGlnLeuValaLeu 96
Db 475 CGCAAGACTACAACTGACAGAGAGAAAGCTTGTGCGTGTGATTAATTGATTCA 534
Qy 97 ValaAspH1aProAsn1IeValaLysLeuH1aLysTyrTrpLeuAspThrSerGluAlaCys 116
Db 535 TTGAGCAATCTTAACATTTGTTAACTTCAAAATTTGGCGTGAATTAAGAAACAAG 594
Qy 117 AlaArgVala1IePhe1IeThraGluThraValaSerArgLysSerLeuLysGlnPheLeuLys 136
Db 595 GCCAGGGTCATTTTATACAGAGATACATGATCTGGAGTCTGAAGCAATTTCTGAAG 654
Qy 137 LysThrLysLysAsnH1aLysAlaMetAsnAlaArgAlaTrpLysArgTrpCysThrGln 156
Db 655 AAGAACCAAAAGAACCAACAGACGATGAATGAAGAAAGCATGGAACGTTGGTGCACACA 714
Qy 157 1IeLeuSerAlaLeuSerPheLeuH1aLysCysSerProPhe1IeH1aGlyAsnLeu 176
Db 715 ATCCCTCTGCGCCCTTAAGCTACCTGCACTCTGTGACCCCCCATCATCGGAACCTG 774
Qy 177 ThrSerAspThr1IePhe1IeGlnH1aSerGlyLeu1IeLys1IeGlySerAlaTrpHis 196
Db 775 ACCTGTGACACCATTTATCCAGCAACGACATCAAGATTGAGCTGTG----- 828
Qy 197 Arg1IePheSerAsnAlaLeuArgProThrAlaLeuProAspAspLeuArgSerPro 216
Db 829 -----GCTCTGACACATCAACATCAT 852
Qy 217 1IeArgAlaGluArgGluGluLeuArgAsnLeuH1aPhePheProProGluTyrGlyGlu 236
Db 853 GTGAAGACTTGGCGAAGAGCAAGAAATCTTACCTTTGACACAGATGATGAGAA 912
Qy 237 Vala1aAsp---GlyThrAlaValaAsp1IePheSerPheGlyMetCysAlaLeuGluMet 255


```

; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS
; TITLE OF INVENTION: OF LUNG CANCER
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,984
; FILING DATE: 18-MAR-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.456
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-622-4800
; TELEFAX: 206-282-6031
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-040-984-23
;
Alignment Scores:
Pred. No.: 8,32e-33 Length: 669
Score: 383.50 Matches: 83
Percent Similarity: 66.24% Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match: 14.36% Indels: 13
Gaps: 3
;
US-10-618-941-99 (1-507) x US-09-040-984-23 (1-669)
Oy 349 AlAValleuAlaGlueuProArG---ProArGArGProProleuGIntPrArGTYrSer 367
Db 7 GCCGACTGGCTGAATCCCTGCAGACCAAGAGAAACAGTTCAAGCTTTGACTCT 66
Oy 368 GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTYrPro 387
Db 67 CAGTCACCAAGCTCTGAATTAGATAATTCCTTGAAGATGTCAGAAATGGGATCTATCT 126
Oy 388 LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValIleuAlaProProBro 407
Db 127 CTGACAGCTTT-----GGGCTGCTCGGCCGCCACAGCCACAGCAG 168
Oy 408 GluGluVal-----GlnLysAlaLysThrProThrProGluuProPhe 421
Db 169 GAGAGGGTGAATCATCTGTCTGTGCCCCCTGTCTCAAGACTCCGACACTGAACCACT 228
Oy 422 AspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLys 441
Db 229 GAGGTGAGAGACTCCGAAAGTGTGCTGATGCAATTCATTCGATCGGTGAGAGAGGA 288
Oy 442 AlaArgTrpHisLeuThrLeuLeuValLeuGluAspArgLysHisArgGlnLeuThr 461
Db 289 GTCAAAACACCACTGACACTTCTCTGAAGTTGAGAGCAAACTGAACCGGACCTGAGC 348
Oy 462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisArgIly 481
```

```

Db 349 TGTGACCTGAGCCAAAGAGAAATATCCCCAGATTGGCGGCTGAGCTGTGACGCTGGC 408
Oy 482 PheLeuHisGluAspAspArgMetLysLeuAlaAlaPheLeuGluSerThr 498
Db 409 TTCTATTAGTGAGCTGACACGAGCGGTTGACTTCTCTGTAGAAGAGACT 459
;
RESULT 11
US-09-123-912-23
; Sequence 23, Application US/09123912A
; Patent No. 6312695
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tonglong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER
; FILE REFERENCE: 210121.455C1
; CURRENT APPLICATION NUMBER: US/09/123,912A
; CURRENT FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: 09/040,802
; PRIOR FILING DATE: 1998-03-18
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 23
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (642)
; OTHER INFORMATION: where n is a, c, g or t
; NAME/KEY: modified_base
; LOCATION: (661)
; OTHER INFORMATION: where n is a, c, g or t
;
US-09-123-912-23
;
Alignment Scores:
Pred. No.: 8,32e-33 Length: 669
Score: 383.50 Matches: 83
Percent Similarity: 66.24% Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match: 14.36% Indels: 13
Gaps: 3
;
US-10-618-941-99 (1-507) x US-09-123-912-23 (1-669)
Oy 349 AlAValleuAlaGlueuProArG---ProArGArGProProleuGIntPrArGTYrSer 367
Db 7 GCCGACTGGCTGAATCCCTGCAGACCAAGAGAAACAGTTCAAGCTTTGACTCT 66
Oy 368 GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTYrPro 387
Db 67 CAGTCACCAAGCTCTGAATTAGATAATTCCTTGAAGATGTCAGAAATGGGATCTATCT 126
Oy 388 LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValIleuAlaProProBro 407
Db 127 CTGACAGCTTT-----GGGCTGCTCGGCCGCCACAGCCACAGCAG 168
Oy 408 GluGluVal-----GlnLysAlaLysThrProThrProGluuProPhe 421
Db 169 GAGAGGGTGAATCATCTGTCTGTGCCCCCTGTCTCAAGACTCCGACACTGAACCACT 228
Oy 422 AspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLys 441
Db 229 GAGGTGAGAGACTCCGAAAGTGTGCTGATGCAATTCATTCGATCGGTGAGAGAGGA 288
Oy 442 AlaArgTrpHisLeuThrLeuLeuValLeuGluAspArgLysHisArgGlnLeuThr 461
Db 289 GTCAAAACACCACTGACACTTCTCTGAAGTTGAGAGCAAACTGAACCGGACCTGAGC 348
Oy 462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisArgIly 481
Db 349 TGTGACCTGAGCCAAAGAGAAATATCCCCAGATTGGCGGCTGAGCTGTGACGCTGGC 408
```

```

Qy      482 PheLHLEGLUAspAArgMetLysLeuAlaPheLeuGlnSerThr 498
Db      409 TTCATTAGTAGGCTGCACGAGCCGGTTGACTTCTCTGTAAGAAGACT 459

RESULT 12
US-09-643-597-23
/ Sequence 23, Application US/09643597
/ Patent No. 6426072
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Fan, Liqun
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Bangur, Chaitanya S.
/ APPLICANT: Hosken, Nancy
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aljun
/ APPLICANT: Skeiky, Yaelir A.W.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: McNeill, Patricia D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C11
/ CURRENT APPLICATION NUMBER: US/09/643,597
/ CURRENT FILING DATE: 2000-08-21
/ NUMBER OF SEQ ID NOS: 369
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 23
/ LENGTH: 669
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(669)
/ OTHER INFORMATION: n = A,T,C or G
US-09-643-597-23

Alignment Scores:
Pred. No.:      8,32e-33      Length:      669
Score:          383.50      Matches:      83
Percent Similarity: 66.24%      Conservative: 21
Best Local Similarity: 52.87%      Mismatches: 40
Query Match:    14.36%      Indels:      13
Db:             3           Gaps:           3

US-10-618-941-99 (1-507) x US-09-643-597-23 (1-669)
Qy      349 AlaValLEuAlaGluLeuProArg--ProArgAArgProProLeuGlnTrpArgTyrSer 367
Db      7 GCCGTACTGGCTGAATCCCTGCAGACGAGACGAGAACAGTTCAGACTTTGTACTCT 66
Qy      368 GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTyrPro 387
Db      67 CAGTCACCAAGCTCTGGAATTGATTAATTCCTTGAAGATGTCAGGATGGATCTATCCT 126
Qy      388 LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValLeuAlaProProPro 407
Db      127 CTGACAGCCTT-----GGGCTGCTCGGCCCCAGACGACGACAGCAG 168
Qy      408 GluGluVal-----GlnLysAlaLysThrProThrProGluProPhe 421
Db      169 GAGGAGGTGACATCACTGTCTGCTGCCCTCTGTCAAGACTCGACACCTTAACCAAGCT 228
Qy      422 AspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLys 441
Db      229 GAGGTGAGACTCGCAAGGTGGTGGTGAAGTGAACATTGAGTCGGTGAAGGAGGA 288
Qy      442 AlaArgTrpHisLeuThrLeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThr 461
Db      289 GTCAAAACACCACTGACCTTCTGCGAAGTGGAGGACAACTGAACCGGACCTGAGC 348
Qy      462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaIlePheLeuGlnSerThr 481

```

```

Db      349 TGTGACTGATGCCAAATGAGATATCCCGAGTTGGCGGCTGAGCTGGTGCAGCTGGGC 408
Qy      482 PheLHLEGLUAspAArgMetLysLeuAlaPheLeuGlnSerThr 498
Db      409 TTCATTAGTAGGCTGCACGAGCCGGTTGACTTCTCTGTAAGAAGACT 459

RESULT 13
US-09-480-884A-23
/ Sequence 23, Application US/09480884A
/ Patent No. 6482597
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Tongtong
/ APPLICANT: Fan, Liqun
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Hosken, Nancy A.
/ APPLICANT: Fanger, Gary R.
/ APPLICANT: Li, Samuel X.
/ APPLICANT: Wang, Aljun
/ APPLICANT: Skeiky, Yaelir A.W.
/ APPLICANT: Henderson, Robert A.
/ APPLICANT: McNeill, Patricia D.
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY
/ TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
/ FILE REFERENCE: 210121.455C6
/ CURRENT APPLICATION NUMBER: US/09/480,884A
/ CURRENT FILING DATE: 2001-08-27
/ NUMBER OF SEQ ID NOS: 330
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 23
/ LENGTH: 669
/ TYPE: DNA
/ ORGANISM: Homo sapien
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(669)
/ OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-23

Alignment Scores:
Pred. No.:      8,32e-33      Length:      669
Score:          383.50      Matches:      83
Percent Similarity: 66.24%      Conservative: 21
Best Local Similarity: 52.87%      Mismatches: 40
Query Match:    14.36%      Indels:      13
Db:             3           Gaps:           3

US-10-618-941-99 (1-507) x US-09-480-884A-23 (1-669)
Qy      349 AlaValLEuAlaGluLeuProArg--ProArgAArgProProLeuGlnTrpArgTyrSer 367
Db      7 GCCGTACTGGCTGAATCCCTGCAGACGAGACGAGAACAGTTCAGACTTTGTACTCT 66
Qy      368 GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTyrPro 387
Db      67 CAGTCACCAAGCTCTGGAATTGATTAATTCCTTGAAGATGTCAGGATGGATCTATCCT 126
Qy      388 LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValLeuAlaProProPro 407
Db      127 CTGACAGCCTT-----GGGCTGCTCGGCCCCAGACGACGACAGCAG 168
Qy      408 GluGluVal-----GlnLysAlaLysThrProThrProGluProPhe 421
Db      169 GAGGAGGTGACATCACTGTCTGCTGCCCTCTGTCAAGACTCGACACCTTAACCAAGCT 228
Qy      422 AspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLys 441
Db      229 GAGGTGAGACTCGCAAGGTGGTGGTGAAGTGAACATTGAGTCGGTGAAGGAGGA 288
Qy      442 AlaArgTrpHisLeuThrLeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThr 461
Db      289 GTCAAAACACCACTGACCTTCTGCGAAGTGGAGGACAACTGAACCGGACCTGAGC 348
Qy      462 TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaIlePheLeuGlnSerThr 481
Db      349 TGTGACTGATGCCAAATGAGATATCCCGAGTTGGCGGCTGAGCTGGTGCAGCTGGGC 408
Qy      482 PheLHLEGLUAspAArgMetLysLeuAlaIlePheLeuGlnSerThr 498

```

```
Db 409 TTCATTAGTGAAGCTGACCAAGCCGGTTGACTTCTGCTAGAAAGACT 459
RESULT 14
US-09-542-615A-23
; Sequence 23, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hoeken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(669)
; OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-23

Alignment Scores:
Pred. No.: 8.32e-33 Length: 669
Score: 383.50 Matches: 83
Percent Similarity: 66.24% Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match: 14.36% Indels: 13
DB: 3 Gaps: 3

US-10-618-941-99 (1-507) x US-09-542-615A-23 (1-669)
Qy 349 A1aVal1euA1aG1u1euProArg---ProArgArgProProleuG1nTrpArgTySer 367
Db 7 GCCGTACGCTGTAATCCCTCGACAGACCAAGAAACCAAGTTCAAGCTTGACTCT 66
Qy 368 GluVal1SerPheMetG1u1euA8rlyPhe1euG1uA8rvalA1aG8nG1y11eTyP8r 387
Db 67 CAGTCACACAGCTCTGAATTAATTAATTCCTGAAGATGTCAGGAATGGGATCTATCT 126
Qy 388 LeuMetAnPheA1a1a1aThrArgProleuG1y1euProArgVal1euA1aProPro 407
Db 127 CTGACAGCCCTT-----GGGCTGCTCGGCCGCCACAGCCACAGCAG 168
Qy 408 GluGluVal1-----G1nlyA1a1y8TrP8rProG1uP8rPhe 421
Db 169 GAGAGAGGACATCATCTGTCTGCCCCCTCTGCAAGATCCGACACCTGAACCACT 228
Qy 422 A8pSerG1uThrArgLy8Val11eG1nMetG1nCy8aen1euG1uA8rSerG1uA8rly8 441
Db 229 GAGGTGAGACTCCGAAGTGTGTGATGCACTGAACATTAAGTCGTGAGAGGGA 288
Qy 442 A1aArgTrpH1e1euThr1eu1euVal1euG1uA8rA8rly8n1a8rG1n1euThr 461
Db 289 GTCAAAACCAACCTGACATTTCTGTGAAGTGGAGCAACATGAACCGGACCTGAGC 348
Qy 462 TyrA8r1eu1euProThrA8pSerA1aG1nA8r1euA1aSerG1u1euVal1H1eTy8Ly 481
Db 349 TGTACCTGATGCCAAATGAATAATCCCGAGTTGGCGGCTGAGCTGTGACGTGGGC 408
Qy 482 Phe1euH1eG1uA8rA8rArgMetLy81euA1a1aPhe1euG1uSerThr 498
Db 409 TTCAATTAGTGAAGCTGACCAAGCCGGTTGACTTCTGCTAGAAAGACT 459
,RESULT 15
```

```
US-09-606-421B-23
; Sequence 23, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hoeken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(669)
; OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-23

Alignment Scores:
Pred. No.: 8.32e-33 Length: 669
Score: 383.50 Matches: 83
Percent Similarity: 66.24% Conservative: 21
Best Local Similarity: 52.87% Mismatches: 40
Query Match: 14.36% Indels: 13
DB: 3 Gaps: 3

US-10-618-941-99 (1-507) x US-09-606-421B-23 (1-669)
```

```
Qy 349 A1aVal1euA1aG1u1euProArg---ProArgArgProProleuG1nTrpArgTySer 367
Db 7 GCCGTACGCTGTAATCCCTCGACAGACCAAGAAACCAAGTTCAAGCTTGACTCT 66
Qy 368 GluVal1SerPheMetG1u1euA8rlyPhe1euG1uA8rvalA1aG8nG1y11eTyP8r 387
Db 67 CAGTCACACAGCTCTGAATTAATTAATTCCTGAAGATGTCAGGAATGGGATCTATCT 126
Qy 388 LeuMetAnPheA1a1a1aThrArgProleuG1y1euProArgVal1euA1aProPro 407
Db 127 CTGACAGCCCTT-----GGGCTGCTCGGCCGCCACAGCCACAGCAG 168
Qy 408 GluGluVal1-----G1nlyA1a1y8TrP8rProG1uP8rPhe 421
Db 169 GAGAGAGGACATCATCTGTCTGCCCCCTCTGCAAGATCCGACACCTGAACCACT 228
Qy 422 A8pSerG1uThrArgLy8Val11eG1nMetG1nCy8aen1euG1uA8rSerG1uA8rly8 441
Db 229 GAGGTGAGACTCCGAAGTGTGTGATGCACTGAACATTAAGTCGTGAGAGGGA 288
Qy 442 A1aArgTrpH1e1euThr1eu1euVal1euG1uA8rA8rly8n1a8rG1n1euThr 461
Db 289 GTCAAAACCAACCTGACATTTCTGTGAAGTGGAGCAACATGAACCGGACCTGAGC 348
Qy 462 TyrA8r1eu1euProThrA8pSerA1aG1nA8r1euA1aSerG1u1euVal1H1eTy8Ly 481
Db 349 TGTACCTGATGCCAAATGAATAATCCCGAGTTGGCGGCTGAGCTGTGACGTGGGC 408
Qy 482 Phe1euH1eG1uA8rA8rArgMetLy81euA1a1aPhe1euG1uSerThr 498
Db 409 TTCAATTAGTGAAGCTGACCAAGCCGGTTGACTTCTGCTAGAAAGACT 459
```

Search completed: January 14, 2006, 00:50:21
Job time : 275 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

; Sequence 99, Application US/10618941
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 99
; LENGTH: 507
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-618-941-99
MAPPPAPRRARRRRERREDESEDSIDLESPQGRWQKRREOVNOGMPQLOSTPIAMDTEEGV
EVWNLHFDGRKAFARHEKIQTVFQVLVDHPNIVKJHKWLDTSKACARVIFTETYSGS
LKQPLKTKKNHAKMARAKRWCTQLLSLSPHACSPRIHGNLTSDTIFIQNGLIKISVM
HRTFSNALPPPTALPDRLRSPIAREBELNLFPPPEYGEVADGTAVDIFSGMCALMAVLEI
QTNGDTRVTEEAJARAHSLSDDPMREFILCCCLADPARPSASHLSLFHRVLFVESHKLLAAHC
FIQHOYLMENNVYEKTKAMDLAVALPRPRPIQWRYSVSEFWELDKPLEDVNGIYPLMN
PAATRPLGLPRVLAAPPPEVQAKTTPPEPDSSTRKYIOMCNLERSEDKARNHLLTLVLVEDR
LHRQLTYDLPTDSADLASELVHYGFLHEDRKLALFLESTFLKTYGTQAI

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2006, 11:16:54 ; Search time 75 Seconds

(without alignments)
4769.371 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAAPBPAPRABRABRERERED.....RMKLAAPLESTFLKRGCTQA 507

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1562.5	58.5	547	2	Q8BL77_MOUSE
2	1541	57.7	526	2	Q6GLJ1_XENTR
3	1537.5	57.6	535	2	Q9J9J5_MOUSE
4	1536	57.5	294	2	Q8NCK8_HUMAN
5	1534.5	57.5	535	2	Q9UHY1_HUMAN
6	1534.5	57.5	535	2	Q4R8X0_MACFA
7	1534.5	57.5	535	2	Q5RBH9_PONPY
8	1529.5	57.3	535	2	Q96SU3_HUMAN
9	1527.5	57.2	535	2	Q5JF25_HUMAN
10	1519	56.9	532	2	Q5ZHL8_CHICK
11	1514	56.7	526	2	Q6NTZ7_XENLA
12	1486.5	55.7	519	2	Q5SNQ3_BRARE
13	1400.5	52.5	501	2	Q4SIG5_TETNG
14	1346	50.4	258	2	Q9NSY0_HUMAN
15	1331.5	49.9	483	2	Q4RKPS_TETNG
16	1283	48.1	258	2	Q9IV36_MOUSE
17	1174.5	44.0	637	2	Q9Y0V6_DROME
18	1170	43.8	293	2	Q8WY55_HUMAN
19	1160	43.4	484	2	Q4T5T2_TETNG
20	1145.5	42.9	425	2	Q7O3R5_ANOGA
21	965	35.6	324	2	Q5Y7G7_HUMAN
22	951.5	35.6	649	2	Q45668_CAEBL
23	778	29.1	157	2	Q8R3M0_MOUSE
24	395	14.8	2193	2	Q5VWF2_HUMAN
25	395	14.8	2297	1	WNK2_HUMAN
26	392	14.7	379	2	Q7PUA4_ANOGA
27	378	14.2	848	2	Q56V74_MOUSE
28	378	14.2	1222	1	WNK4_MOUSE
29	378	14.2	1222	2	Q4VAC1_MOUSE
30	378	14.2	1677	2	Q8I127_CAEBL
31	378	14.2	1743	1	WNK3_HUMAN

32	378	14.2	1743	2	Q5JRC1_HUMAN	Q5JRC1_homo sapien
33	378	14.2	1790	2	Q5JRC2_HUMAN	Q5JRC2_homo sapien
34	378	14.2	1800	2	Q8TCK6_HUMAN	Q8TCK6_homo sapien
35	378	14.2	1800	2	Q6UP76_HUMAN	Q6UP76_homo sapien
36	378	14.2	1838	2	Q18657_CAEBL	Q18657_caenorhabdi
37	377.5	14.1	438	2	Q53PL3_ORYZA	Q53PL3_oryza sativ
38	376	14.1	439	2	Q6DCU2_XENLA	Q6DCU2_xenopus lae
39	372	13.9	516	2	Q9STK6_ARATH	Q9STK6_arabidopsis
40	372	13.9	1222	1	WNK4_RAT	Q7PK6_rattus norv
41	370.5	13.9	1243	1	WNK4_HUMAN	Q96J92_homo sapien
42	370	13.9	1666	2	Q60QJ1_CAEBR	Q60QJ1_caenorhabdi
43	368.5	13.8	568	2	Q8S8Y9_ARATH	Q8S8Y9_arabidopsis
44	366	13.7	578	2	Q9LUV8_ARATH	Q9LUV8_arabidopsis
45	363.5	13.6	600	2	Q8H4Y8_ORYZA	Q8H4Y8_oryza sativ

ALIGNMENTS

RESULT 1
Q8BL77_MOUSE PRELIMINARY; PRT; 547 AA.
ID Q8BL77_MOUSE
AC Q8BL77_MOUSE
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length
DE enriched library, clone:B2034417 product:SIMILAR TO NUCLEAR RECEPTOR
DE BINDING PROTEIN (HLS7-INTERACTING PROTEIN KINASE) homolog.
GN Name=Nrbp;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Eularchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J; TISSUE=Corpora quadrigemina;
RA MEDLINE=92279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[2]
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J; TISSUE=Corpora quadrigemina;
RA MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shinaigawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant H.,
RA Frieschmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirral L.M., Straubli P., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Orido T., Furuno M., Hono H., Baldarelli R., Barin G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohatsu S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RN
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CS7BL/6J; TISSUE=Corpora quadrigemina;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:11617-1630(2000).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama Y., Nishi K., Katsunai T., Taenhiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Yoneda Y., Ishikawa T., Ozawa K., Kira A., Hayashizaki Y.;
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J; TISSUE=Corpora quadrigemina;
RA Adachi U., Aizawa K., Akimura T., Arikawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Horii F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Kono H., Kurauchi S.,
RA Kurahara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takeuchi-Akita S., Takeda Y., Tanaka T.,
RA Tomaru A., Taya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK046142; BAC32612.1; -; mRNA.
DR Ensembl; ENSMUSG00000029148; Mus musculus.
DR MGI; MGI:2183436; Nrdp.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR Pfam; PF00069; Pkinase; 1.
DR Prodom; PD000001; Prot kinase; 1.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
DR ATP-binding; Kinase; Nucleotide-binding; Receptor; Transferase.
SQ SEQUENCE 547 AA; 61221 MW; 57462EA936E2762 CRC64;
Qy Query Match 58.5%; Score 1562.5; DB 2; Length 547;
Best Local Similarity 62.6%; Pred. No. 2.7e-104;
Matches 310; Conservative 60; Mismatches 104; Indels 21; Gaps 6;
Qy 17 EREDESEDESDILSPGCRGWRQKREOVNGNMPGLQSTFLAMDTEGVEVNNELAFGD 76
Db EEESESEDESEILSPGCRGWRQKREOVNGNMPGLQSTFLAMDTEGVEVNNELAFGD 103
Qy 77 RKAPFAAHEKIQVFEQVLVVDHPNIVYKLYWLDLSEACARVITFTTEVSSGLKQFLK 136
Db RKNYKIQOEKVRVAFNLIQLEHNTIVKFKYVADWENKARVITFTTEVSSGLKQFLK 163
Qy 137 KTKGNKHAMNARAKMCKCTOILSALSIHACSPPIIHGNTSTPIFHNGLIKISWVH 196
Db KTKGNKHAMNARAKMCKCTOILSALSIHACSPPIIHGNTSTPIFHNGLIKISWVH 223
Qy 197 RIFSNALRPETALPDDLSPIRAERBELNIAHFPPEYGEVAD-GTAVDIFSGMCALEM 255
Db RIFSNALRPETALPDDLSPIRAERBELNIAHFPPEYGEVAD-GTAVDIFSGMCALEM 277
Qy 224 RIFAN-----VAPDITNNHVKTGREQKQLHFFABBYGSEVNTTAVDLYSGMCALEM 277

Qy 256 AVLEIOTNGDTR-VTEBATAARHSLSDPMNEPILCCARDPARPSASHLFFHVLPE 314
Db AVLEIOTNGDESSVPEALISSAIIQLLEDSLQREFFIQKQSQSPARPTARELLFHPALPE 337
Qy 315 VHSILKLAACHACTIOHQYLMPEVNEVEEKTAMPLHVALELPR-PRRPPQWRSEVSFMS 373
Db VHSILKLAACHACTIOHQYLMPEVNEVEEKTAMPLHVALELPR-PRRPPQWRSEVSFMS 397
Qy 374 LDKPLEDVANGIYPLMNFATRPGLPRVLADEPPEV-----QKAKTPPEPDSERTR 427
Db LDKPLEDVANGIYPLMNFATRPGLPRVLADEPPEV-----QKAKTPPEPDSERTR 451
Qy 428 VYQMCNLERSDKARKMTLLVLVEDRLHRLQTLDTLPTDQAQDLASLVNHFLEHED 487
Db VYQMCNLERSDKARKMTLLVLVEDRLHRLQTLDTLPTDQAQDLASLVNHFLEHED 511
Qy 488 RKMALAFESTFLKY 502
Db RKMALAFESTFLKY 526
Qy 512 QSRILSVLEETLANK 526
Db QSRILSVLEETLANK 526
RESULT 2
Q6GLJ1 XENTR
ID Q6GLJ1 XENTR PRELIMINARY; PRT; 526 AA.
AC Q6GLJ1
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Nuclear receptor binding protein.
GN Name=nrp-prov;
OC Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus; Silurana.
NCBI_TaxID=8364;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustun T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butlerfield Y.S.N., Krzywicki M.T., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Maitra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Embryo;
RX Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC074684; AAH74684.1; -; mRNA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0004872; F:protein kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_kinase.


```

Db      164 KTKGNKHTNKKMKMKWCIOILSALSTHSQCDPIIHGNTCTTITQNHGLIKISV-- 221
QY      197 RIFSNALRPPTALPDILRSPIRAERBELRNLFPPPEYGEVAD-GRVADIFSGWCALEM 255
Db      222 -----APDTINNHVKTCREQKALHFAPEYGEVTVTTAVDVISFGWCALEM 269
QY      256 AVLEIQTNQDTR-VTEBAIARAHSLSDPMWREFIICCLARDARRBSASLLEFHVLEP 314
Db      270 AVLEIQNGESSYVPOALSSAIQLLEDSTIQREFIQKQSEPARPPTAEELLFHALPE 329
QY      315 VHSIKLLAAHCFIIOHLYMPENVEEKTAMDLAAVLALPR-PRRPPLOMRSEVSFME 373
Db      330 VPSLKLLAAHCFIIGHQIMTEPNALLETITKMDTSAVLAETPAPGRBPVQTYLSQSPALE 389
QY      374 LDKFLEDVNRNGIYPLNFAATRPGLPRVLAPPEEV-----QAKTPTPEPDSERK 427
Db      390 LDKFLEDVNRNGIYPLFAF-----GLPRPQOQOEVTSPVPPSVKTPPEPAVEYTRK 443
QY      428 VIOMQCNLERSSEDKARHMLTLLVLEDRRLRQITVDLLPTDSAQDLASELVHYGFLHEDD 487
Db      444 VVLMQCNIESVEEGVKHMLTLTKLEBDKLRHLSCDLMPEISIPDLAELVQGFISEAD 503
QY      488 RMKLAAPLESTFLKRYGTQ 506
Db      504 QSRLTSLLEETLNKFNFTK 522

```

```

RESULT 4
Q8NCX8_HUMAN
ID Q8NCX8_HUMAN PRELIMINARY; PRT; 294 AA.
AC Q8NCX8;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein DKFZp434i2411 (Fragment).
GN Name=NRBP2; Synonyms=DKFZp434i2411;
OS Homo sapiens (human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
NCBI_Taxid=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RG The German cDNA Consortium;
RA Bahr A., Jander J., Mewes H.W., Weil B., Amid C., Oeanger A., Fodo G.,
RA Hen M., Wiemann S.;
RA Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DB EMBL; ALB34530; CAD39186.1; -; mRNA.
DB Ensembl; ENSG00000185189; Homo sapiens.
DB HGNC; HGNC:19339; NRBP2.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR ProDom; PD000001; Prot_kinase.
DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
KM Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 294 AA; 33958 MW; DQF57F1901AFB18 CRC64;

Query Match 57.5%; Score 1536; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 9,7e-103;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      334 PENVVEEKTAMDLAAVLALPRRPRPPLQWYSEVSFMEIDKFLQEDVNRNGIYPLNFEA 393
Db      121 PENVVEEKTAMDLAAVLALPRRPRPPLQWYSEVSFMEIDKFLQEDVNRNGIYPLNFEA 180
QY      394 TRPLGLPRVLAPPEEVQKAKTPTEPPEPDSERTRKYOMQCNLERSSEDKARHMLTLLVLE 453
Db      181 TRPLGLPRVLAPPEEVQKAKTPTEPPEPDSERTRKYOMQCNLERSSEDKARHMLTLLVLE 240
QY      454 DLRLHQITVDLLPTDSAQDLASELVHYGFLHEDDNRKLAAPLESTFLKRYGTQ 507
Db      241 DLRLHQITVDLLPTDSAQDLASELVHYGFLHEDDNRKLAAPLESTFLKRYGTQ 294

RESULT 5
Q9UHY1_HUMAN
ID Q9UHY1_HUMAN PRELIMINARY; PRT; 535 AA.
AC Q9UHY1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 13-MAY-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Multiple domain putative nuclear protein (Hypothetical protein
DE Ffj11084) (Hypothetical protein DKFZp564D1878) (Nuclear receptor
DE binding protein) (Hypothetical protein NRBP).
GN Name=BCON3; Synonyms=DKFZp564D1878, NRBP;
OS Homo sapiens (human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae;
OC Homo.
NCBI_Taxid=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20304761; PubMed=10843813; DOI=10.1006/geno.2000.6167;
RA Hooper J.D., Baker E., Ogbourne S.M., Sutherland G.R., Antalis T.M.;
RT "Cloning of the cDNA and localization of the gene encoding human NRBP,
RT a ubiquitously expressed, multidomain putative adapter protein.";
RL Genomics 66:113-118(2000).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Nakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obedyashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuna M.,
RA Shitatori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishidaishi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirooka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Toguchi S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuki H., Oshima A., Saeki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Omori Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Saeki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita K.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RN Nat. Genet. 36:40-45(2004).

```

RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Brain;
 RG The German cDNA Consortium;
 RA Oettermeyer B., Obermeier B., Deutschenaur S., Schapp A.,
 RA Mewes H.W., Weill B., Amlid C., Oeinger A., Fodor G., Han M., Wiemann S.;
 RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Lung;
 MDLINE=22368257; PubMed=12477932; DOI=10.1073/pnas.242603699;
 RA Strauberg R.L., Feligold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diachenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Urdin T.B., Tothlyuk S., Carninci P., Prange C.,
 RA Bosa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Bailey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smalhe D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Matra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE-Lung;
 RA Strauberg R.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Sun H., Kozlowicz A., Dignan G.;
 RT "The sequence of Homo sapiens BAC clone RP11-113M20.";
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.H.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Waterston R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Wilson R.K.;
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF113249; AAF21967.1; -; mRNA.
 DR EMBL; AK001946; BAA91993.1; -; mRNA.
 DR EMBL; AL136682; CAB66617.1; -; mRNA.
 DR EMBL; BC001221; AA01221.1; -; mRNA.
 DR EMBL; AC074117; AAT14847.1; -; Genomic DNA.
 DR Ensembl; ENSG00000115216; Homo sapiens.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0005070; F:SH3/SH2 adaptor activity; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR InterPro; IPR000719; Prot_kinase.
 DR Pfam; PF00069; Kinase; 1.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR POSITIVE; PSS0011; PROTEIN KINASE DOM; 1.
 KM ATP-binding; Hypothetical protein; Kinase; Nuclear protein;
 KM Nucleotide-binding; Receptor; Serine/threonine-protein kinase;
 KW Transference; 535 AA; 59845 MW; 398078661547EDD0 CRC64;
 SQ SEQUENCE

QY 17 EREDESESDILEESPGRGKRRROVNGNMPGLSTFLAMDTSEGVVNNELHFGD 76
 DB 44 EEEESSEDESEILEESPGRGKRRROVNGNMPGLSTFLAMDTSEGVVNNELHFGD 103
 QY 77 RKAFAAHEEKIQTFEQLVLDHPNIVKLRWLDTSSEACARVIFITEVSSGLKQFLK 136
 DB 104 RKNYKLOEKEVRAVFDNLQLEHNTVFKFKYVADIKENKARVIFITEVSSGLKQFLK 163
 QY 137 KTKGNKAMAKAMKWCCTOILSALSFACSPPIIHGNTISPTIFQNHGLIKGSVMH 136
 DB 164 KTKGNKAMAKAMKWCCTOILSALSYLHSCDPIIHGNTICPTIFQNHGLIKGSVMH 221
 QY 197 RIFSNMLRPTALPDRLSPRABREIRLWHFPPEGVAD-GRVADIFSGMCALM 255
 DB 222 -----APDTINNVKTCREBQKMLHFPPEGVADIFSGMCALM 269
 QY 256 AVLEIQNGDTR-VTEBAIARASHSLDPNMRFFICLARDPARPSAHLFHRVLP 314
 DB 270 AVLEIQNGESSVVPQEAISALQLEDPQREFIKCQSEPARPTARELLFHPALFE 329
 QY 315 VHSKLIAAHCFIQHQTLPENVEEKTAMDIAVLAELPR-PRRPLQWRYSSEVME 373
 DB 330 VPSLKLIAAHCFIQHQTLPENVEEKTAMDIAVLAELPR-PRRPLQWRYSSEVME 389
 QY 374 LDKPLDVRNGIYPLMNPATRLGLPRVLAPEPEE-----QKATPPEPDSSTRK 427
 DB 390 LDKPLDVRNGIYPLMNPATRLGLPRVLAPEPEE-----QKATPPEPDSSTRK 443
 QY 428 VIOMQNLERSSEDKARWHLTLVLVLEDRLRQTYVLDPTDSAQDLASELVHGFLEDD 487
 DB 444 VLMQCNIBSEBEGVGHLLTLKLEKLNRLHSCDLMENRIPBLAELVQGFISEAD 503
 QY 488 RMTLAAPLESTFLKY 502
 DB 504 QSRLLTSLEBETLNKF 518

RESULT 6
 Q4R8X0_MACPA
 ID Q4R8X0_MACPA PRELIMINARY; PRT; 535 AA.
 AC Q4R8X0;
 DT 13-SEP-2005 (T-EMBLrel. 31, last sequence update)
 DT 13-SEP-2005 (T-EMBLrel. 31, last annotation update)
 DE Testis cDNA clone: Q4R8-11263, similar to human nuclear receptor
 binding protein (NRBP), (Brain cDNA, clone: Q4R8-17237, similar to
 human nuclear receptor binding protein (NRBP)).
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 OC Cercopithecoidea; Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA International consortium for macaque cDNA sequencing, analysis;
 RT "DNA sequences of macaque genes expressed in brain or testis and its
 RT evolutionary implications.";
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 RN
 RP NUCLEOTIDE SEQUENCE.
 RA Osada N., Hirata M., Tanuma R., Kuehda J., Hida M., Suzuki Y.,
 RA Sugano S., Gojobori T., Shen J.C.-K., Wu C.I., Hashimoto K.;
 RT "Substitution rate and structural divergence of 5'UTR evolution:
 RT comparative analysis between human and cynomolgus monkey cDNAs.";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB168327; BAB0451.1; -; mRNA.
 DR EMBL; AB169835; BAB0451.1; -; mRNA.
 KW Receptor.
 KW SEQUENCE 535 AA; 59845 MW; 398078661547EDD0 CRC64;
 SQ SEQUENCE

Query Match 57.5%; Score 1534.5; DB 2; Length 535;
 Best Local Similarity 61.6%; Pred. No. 2.7e-102;
 Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

Matches	305;	Conservative	59;	Mismatches	102;	Indels	29;	Gaps	6
QY	17	EREDSEDESDILIESPCGRMQRREBOVNOGMPGLOSTFLAMDREGEVVMNELLHG	76						
Db	44	EEEDSEDESEILIESPCGRMQRREBOVNOGVNVPIDSAVYLMADREGEVVMNELLHG	103						
QY	77	RKAPFAHEKIOTVEQVLVDHNPVVKLHKWVLDTSEACAVITITTEYVSGSLKQFLK	136						
Db	104	RKNYKLOEKKVAVPDNLQLEHNLVVKPHKRWADIKERKAVITITTEYVSGSLKQFLK	163						
QY	137	KTKGNHKAMNAAKMRCTQQLSALSFLLACSPPIIHGNTSDTIFIOHNGIKTGSYH	196						
Db	164	KTKGNHKAMNAAKMRCTQQLSALSYLHSCDPIIHGNTSDTIFIOHNGIKTGSY--	221						
QY	197	RIFSNAIRPPTALPDLLSPRIAREBELNTHFFPPPEYGEVAD-STAVDIFSGMCALEM	255						
Db	222	-----APDTINNHYVTCGEEQGNLHFAFPAEYGEVYNVTAVDIYSFGMCALEM	269						
QY	236	AVLEIQTNQDTR-VTEBAIARARHSISDPNMEFFLCCIARDPARPSAHSILFFRVLFE	314						
Db	270	AVLEIQNGESSYVQGEAISSAIOLLEDPLOREFIOKCIQSSPARPPTARELLFFHPALFE	329						
QY	315	VHSLKLLAACHCIOHOYLMPEVNVVEKTKAMDLHVLAEIPLR-PRRPLOMYSEVSIME	373						
Db	330	VPSLKLTLAAHCIVGHQHMIPENALBETITQMDTSAVLAIEIPAGPREPEVQTLTSSOPALE	389						
QY	374	LDKFLPDYVNGIYPLMNFATATPLGLPVLADPPPEV-----QKAKTPTPEPDSERTK	427						
Db	390	LDKFLPDYVNGIYPLTAF-----GLPRQOQOGEVYSPVPVPSVKPTTPPEPAVERRK	443						
QY	428	VIQMCNLERSEDKARWHLTLVLVEDRLHROLTYDLLPTDSAQDLASLVHYGFLHED	487						
Db	444	VLMQCNIESVEEGVGHLLTLTKLEDKILRLHSCDLPMBNENIPELAAELVQLGFISEAD	503						
QY	488	RMKLAAPLESTFLKY 502							
Db	504	QSLRTLSLETLNKP 518							
RESULT 7									
OSRBH9_PONPY									
AC	Q5RBH9;	PONPY PRELIMINARY;	PRT;	535	AA.				
DT	01-FEB-2005	(TREMBLrel. 29, Created)							
DT	01-FEB-2005	(TREMBLrel. 29, Last sequence update)							
DT	01-FEB-2005	(TREMBLrel. 29, Last annotation update)							
DE	Hypothetical protein DKFZp459G2032.								
GN	Name=DKFZp459G2032;								
OS	Pongo pygmaeus (orangutan).								
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;								
OC	Pongo.								
OX	NCBI_Taxid=9600;								
RN	(1)								
RP	NUCLEOTIDE SEQUENCE.								
RC	TISUE=Context;								
RA	The German cDNA Consortium;								
RA	Koehler K., Beyer A., Wiemann S., Mewes H.W., Well B., Amid C., Oanger A.,								
RL	Froh G., Han M., Wiemann S.,								
RL	Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.								
DR	EMBL, CR858669; CAH90881.1; -, mRNA.								
DR	GO: GO:0005524; F:ATP binding; IEA.								
DR	GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.								
DR	GO: GO:0004713; F:protein-tyrosine kinase activity; IEA.								
DR	GO: GO:0016740; F:transferase activity; IEA.								

DR	PROSITE, PSS0011; PROTEIN_KINASE_DOM; 1.
KW	Hypothetical protein.
SD	SEQUENCE 535 AA; 59818 MW; 398078660FEEDD0 CRC64;
Query Match	57.5%; Score 1534.5; DB 2; Length 535;
Best Local Similarity	61.6%; Pred. No. 2.7e-102;
Matches 305; Conservative	59; Mismatches 102; Indels 29; Gaps 6;
QY	17 ERDESEDSIDLEESPCGRWOKRREOVNGNPGLOSTFLAMDTEEGVEVNNNELHFGD 76
DB	44 EEESESDSESLLEESPCGRWQRRREVNQNRVPGIDSLAMDTEEGVEVNNNEVQFSE 103
QY	77 RKAFPAHEKIQVFEQVLVDHFNIVYKHKYMLDTSSEACARVIFTEYVSSGLQFLK 136
DB	104 RKNYKQBEKRAVFDMLIOLEHNLIVYKFKYVADIKENKARVIFTEYVSSGLQFLK 163
QY	137 KTKGNHAKMABAKRWCCTQLSLASTLHACSPPIIHGNTLSTPIFHNGGLIKGSVMH 196
DB	164 KTKGNHAKMABAKRWCCTQLSLASTLHACSPPIIHGNTLSTPIFHNGGLIKGSVMH 221
QY	197 RIFSNALRPPTALPDDLRSPIDRAREEELRYMLHFPPEYGEVAD-GTAVDIFSGMCALYM 255
DB	222 -----ADPTINNHVKCTREGRQKRLHFPPEYGEVNTVTTAVDLYSGMCALYM 269
QY	256 AVLEIQNGDTR-VTEBALRAPHSLSDPNMREPIICCLARDPARPSAHSLLFHRVLE 314
DB	270 AVLEIQNGGESSYVPGPAISAIQLLEDPLQREFFIQKCLQSEPARREPTARELLFHPALFE 329
QY	315 VHSIKLLAANCFTIOHQLMPENVEEETKAMDHLAVLAELPR-PRRPRLOMRKSEVSEFME 373
DB	330 VPSIKLLAANCFTIOHQMIPENALBEETIKMMDTSAVLAELIPAGRGREPVQTLVYQSPALE 389
QY	374 LDKLEVDVRNGIYVPLNMFATRPGLGPRVLAAPPEEV-----QKATPPRPDSESTRK 427
DB	390 LDKLEVDVRNGIYVLTAF-----GLPRPOQPOEEVTSVVPSPVTKTPRPEPAEVSTRK 443
QY	428 VIWQCNLERSBDKARWHLTLVLVEDRLRRLQTLTYDLLPTDSADQLASBLVHYGFLHED 487
DB	444 VLVQCNIESVEEGKHLTLTLKLEDKLNRLHSLCDLMPENIPPELAABELVQGLFISEAD 503
QY	488 RMKLAALFLESTFLKY 502
DB	504 QSRLLSLLEETLNKF 518
RESULT 8	
ID	Q96SU3 HUMAN PRELIMINARY; PRT; 535 AA.
AC	Q96SU3; 01-DEC-2001 (TREMBLrel. 19, Created)
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE	Hypothetical protein FLJ14632.
GN	Name=NRBP;
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC	Homo
OX	NCBI_TaxID=9606;
RN	NR
RP	NR
RA	NIHBLORTIDE SEQUENCE.
RA	Isoeqat T., Oca T., Hayaeshi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA	Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,
RA	Wagatsuma M., Hoshiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA	Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,
RA	Wakanebe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,
RA	Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Maehuo Y.,
RA	Nimomiya K., Iwayanagi T.;
RL	Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC	-I- INTERACTION:
CC	PI4340:- (xeno); NBExp=1; IntAct=EBI-465766, EBI-465733;
CC	EMBL; AK027538; BAB55185.1; -, mRNA.
DR	IntAct; Q96SU3; -

Ensembl; ENSG00000115216; Homo sapiens.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR00719; Prot_Kinase.
 DR Pfam; PF00069; Kinase; 1.
 DR Prodom; PD000001; Prot_Kinase; 1.
 DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.
 DR ATP-binding; Kinase; Nucleotide-Binding.
 KW Serine/threonine-protein kinase; transferase.
 KW SEQUENCE 535 AA; 59850 MW; 4829CA7CB60F0175 CRC64;
 Query Match 57.3%; Score 1529.5; DB 2; Length 535;
 Best Local Similarity 61.4%; Pred. No. 6.3e-102;
 Matches 304; Conservative 59; Mismatches 103; Indels 29; Gaps 6;
 QY 17 EREDESDSDILESPGCRKQKREOVNQGMPGLSTFLAMDTEEGVEVNNELHFGD 76
 D 44 EEEESDESEILIESPCGRKQKREOVNQGMPGLSTFLAMDTEEGVEVNNELHFGD 103
 QY 77 RKAFAAEKIQVFEQVLVDHPIYKAKYMLDTSACARVIFTEYVSSGLKQFLK 136
 D 104 RKNYKQEBEKRAVFDNLIOLEHLNIVKFKYMWADIKENKARVIFITEYVSSGLKQFLK 163
 QY 137 KTKKNHAKAMARAKMCTOILSALSPHACSPPIHGNLTSDTIFIOHNGLIKISV 196
 D 164 KTKKNHAKMAMKAKMCTOILSALSYHSCDPIIHGNLTCDTIFIOHNGLIKISV-- 221
 QY 197 RIFSNALRPPTALPDDLSPIRABERELRMHFPPEGEVAD--GTAVIDFSGMCALM 255
 D 222 -----ADPTINNHYKTCREBQKMLHFPAPGEVNTVTTAVDVISFGMCALM 269
 QY 256 AVLEIQNGDTR-VTEBAIARHSLSDPNMRSEFILLCLADPARPSASLHFRVLF 314
 D 270 AVLEIQNGESSYVQGEAISAIQLLEDPLQREFIQCLQSEPARRPTARELLFHPALFE 329
 QY 315 VHSIKLAAHCFIOHVIAMPENVEEKTAMDHLAVLAELPR--PRRPLQMRYSSEV 373
 D 330 VPSIKLAAHCFIOHVIAMPENVEEKTAMDHLAVLAELPR--PRRPLQMRYSSEV 389
 QY 374 LDKLEEDVRNGIYPLMNFATRRPLGLPRVLAAPPEEV-----OKATPTPEPDSSTRK 427
 D 390 LDKLEEDVRNGIYPLTAF-----GLPRPQPOQEEVTSVPVPSVKTTPPEPAVEVTRK 443
 QY 428 VIQMCNLESSEDKARHMLTLVLVEDRLRQLTYYDLPTDSADOLASLAVHGFLEDD 487
 D 444 VVLMQCNIESVEGVKHLTLTLKLEDKLSRHLSCDLMPENIPELAAELVQGLFISEAD 503
 QY 488 RMLKLAFLSTFLKY 502
 D 504 QSRLTSLLEETLTKF 518
 RESULT 9
 053F25 HUMAN PRELIMINARY; PRT; 535 AA.
 AC 053F25;
 DT 13-SEP-2005 (Tremblrel. 31, Created)
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
 DE Nuclear receptor binding protein variant (Fragment).
 GN Name=NRBP;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RA Maruyama K., Sugano S.,
 RT "Oligo-capping: a simple method to replace the cap structure of

RT eucaryotic mRNAs with oligoribonucleotides.";
 RL Gene 138:171-174 (1994).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RA Suzuki Y., Yoshitomo K., Maruyama K., Sugama A., Sugano S.,
 RT "Construction and characterization of a full length-enriched and a 5'-
 RT end-enriched cDNA library."
 RL Gene 200:149-156 (1997).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC Tissue=Kidney;
 RA Suzuki Y., Sugano S., Torok Y., Toyoda A., Takeda T., Sakaki Y.,
 RA Tanaka A., Yokoyama S.,
 RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK223136; BAD96856.1; -; mRNA.
 DR HGNC; HGNC:7993; NRBP.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 KW Receptor.
 KW NON TER
 KW SEQUENCE 535 AA; 59745 MW; 7FD909B54C33E89 CRC64;
 Query Match 57.2%; Score 1527.5; DB 2; Length 535;
 Best Local Similarity 61.4%; Pred. No. 8.8e-102;
 Matches 304; Conservative 59; Mismatches 103; Indels 29; Gaps 6;
 QY 17 EREDESDSDILESPGCRKQKREOVNQGMPGLSTFLAMDTEEGVEVNNELHFGD 76
 D 44 EEEESDESEILIESPCGRKQKREOVNQGMPGLSTFLAMDTEEGVEVNNELHFGD 103
 QY 77 RKAFAAEKIQVFEQVLVDHPIYKAKYMLDTSACARVIFTEYVSSGLKQFLK 136
 D 104 RKNYKQEBEKRAVFDNLIOLEHLNIVKFKYMWADIKENKARVIFITEYVSSGLKQFLK 163
 QY 137 KTKKNHAKAMARAKMCTOILSALSPHACSPPIHGNLTSDTIFIOHNGLIKISV 196
 D 164 KTKKNHAKMAMKAKMCTOILSALSYHSCDPIIHGNLTCDTIFIOHNGLIKISV-- 221
 QY 197 RIFSNALRPPTALPDDLSPIRABERELRMHFPPEGEVAD--GTAVIDFSGMCALM 255
 D 222 -----ADPTINNHYKTCREBQKMLHFPAPGEVNTVTTAVDVISFGMCALM 269
 QY 256 AVLEIQNGDTR-VTEBAIARHSLSDPNMRSEFILLCLADPARPSASLHFRVLF 314
 D 270 AVLEIQNGESSYVQGEAISAIQLLEDPLQREFIQCLQSEPARRPTARELLFHPALFE 329
 QY 315 VHSIKLAAHCFIOHVIAMPENVEEKTAMDHLAVLAELPR--PRRPLQMRYSSEV 373
 D 330 VPSIKLAAHCFIOHVIAMPENVEEKTAMDHLAVLAELPR--PRRPLQMRYSSEV 389
 QY 374 LDKLEEDVRNGIYPLMNFATRRPLGLPRVLAAPPEEV-----OKATPTPEPDSSTRK 427
 D 390 LDKLEEDVRNGIYPLTAF-----GLPRPQPOQEEVTSVPVPSVKTTPPEPAVEVTRK 443
 QY 428 VIQMCNLESSEDKARHMLTLVLVEDRLRQLTYYDLPTDSADOLASLAVHGFLEDD 487
 D 444 VVLMQCNIESVEGVKHLTLTLKLEDKLSRHLSCDLMPENIPELAAELVQGLFISEAD 503
 QY 488 RMLKLAFLSTFLKY 502
 D 504 QSRLTSLLEETLTKF 518
 RESULT 10
 05ZHL8 CHICK PRELIMINARY; PRT; 532 AA.
 AC 05ZHL8;
 DT 25-OCT-2004 (Tremblrel. 28, Created)
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)

```

DT      25-OCT-2004 (TREMblrel. 28, last annotation update)
DE      Hypothetical protein.
GN      ORNames=RCUM004_3518;
OS      Gallus gallus (Chicken).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC      Gallus.
LN      NCBI_TaxId=9031;
RN      [1]
RP      NUCLEOTIDE SEQUENCE.
RC      STRAIN=CB; TISSUE=Bursa;
RA      Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA      Fiedler P., Kutter S., Biagodaeski A., Kostowska D., Koter M.,
RA      Plachy J., Caminci P., Hayashizaki Y., Buerstedt J.M.;
RT      "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT      genefunction analysis.";
RL      Genome Biol. 6:R6-R6(2005).
DR      EMBL; AJ271116; CAG32775.1; -; mRNA.
DR      GO; GO:0005524; F:ATP binding; IEA.
DR      GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR      GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR      GO; GO:0016740; F:transferase activity; IEA.
DR      GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR      InterPro; IPR000719; Prot_Kinase.
DR      InterPro; IPR002290; Ser_thr_kinase.
DR      Pfam; PF00069; PKinase; 1.
DR      Prodom; PD000001; Prot_kinase; 1.
DR      SMART; SM00220; S_TKc.1.
DR      SMART; SM00219; TYRc.1.
DR      ProSITE; PS50011; PROTEIN_KINASE_DOM, 1.
DR      Hypothetical protein.
KW      SEQUENCE 532 AA; 59693 MW; 2176262D10AEC576 CRC64;

```

Query Match	56.9%	Score 151.9;	DB 2;	Length 532;
Beet Local Similarity	59.6%;	Pred. No. 3.6e-101;		
Matches 300;	Conservative 63;	Mismatches 124;	Indels 16;	Gaps 3
QY	6	PAPRAREREREDSESDILESPCGRQKREQVQGNMPLQSTFLAMDTBEGV	65	
DB	33	PYTSAAAPPEEEEESESDSEETLESPCGRQKREQVQGNMPLQSTFLAMDTBEGV	92	
QY	66	EYVNMELHFGDKARPAHEBKIQVFEQVLVDHPNIVKLAKYVLDTSACARVIFTEY	125	
DB	93	EYVNMELHFGDKARPAHEBKIQVFEQVLVDHPNIVKLAKYVLDTSACARVIFTEY	152	
QY	126	VSSGSLKQELKTKTKNKNKAMNARKMWTQTLMSFLHASPRIHENTLSDTFLIOH	185	
DB	153	MSSGSLKQELKTKTKNKNKAMNARKMWTQTLMSFLHASPRIHENTLSDTFLIOH	212	
QY	186	NGLIIGSYWHRIFSNALRPPLPDDLSPRIAREEELRNHFFPEYGEVAD-GTAVD	244	
DB	213	NGLIIGSYWHRIFSNALRPPLPDDLSPRIAREEELRNHFFPEYGEVAD-GTAVD	258	
QY	245	IFSFGMCALMAVLEIQNGDTR-VTEAIALARHSLSDPNKREFILCCLDAPARPSA	303	
DB	259	IYSPGMCALMAVLEIQNGESSYVPOEAINSAIQLEBPLQREVLYQCTBDDPGKRPYA	318	
QY	304	HSLLHFRVLFVEVHSLKTLAAHCFIQOYIMPEVNVBEKTKAMDIAVLAELEFRPPRPQ	363	
DB	319	RELLHFRVLFVEVHSLKTLAAHCFIQOYIMPEVNVBEKTKAMDIAVLAELEFRPPRPQ	378	
QY	364	WRYSEVSFMELEKFLPEDVANGIYPLMNFAATRPGLBRLAPARPEEVQAKTPTDEPPDS	423	
DB	379	WRYSEVSFMELEKFLPEDVANGIYPLMNFAATRPGLBRLAPARPEEVQAKTPTDEPPDS	438	
QY	424	ETRKVIQONCNLSESEDKARMHLLTLVLVEDRLHROLTYDLLPTDSAQDIASELVHYGFL	483	
DB	439	ETRKVIQONCNLSESEDKARMHLLTLVLVEDRLHROLTYDLLPTDSAQDIASELVHYGFL	498	
QY	484	HEDDRMKLAFLLESTFLKTYRGTO 506		
DB	499	SEADQSRLTCLLEBAFSKYYTR 521		

RESULT 11

ID	Q6NT27_XENLA	PRELIMINARY;	PRT;	526 AA.
Q6NT27_XENLA	Q6NT27_XENLA	PRELIMINARY;	PRT;	526 AA.
AC	Q6NT27_XENLA	PRELIMINARY;	PRT;	526 AA.
DT	05-JUL-2004	(TrEMBLrel. 27, Created)		
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)		
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)		
DE	MG681395 protein.			
OS	Name=MG681395;			
OC	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoae; Pipidae;			
OC	Xenopodinae; Xenopus; Xenopus.			
OX	NCBI_Taxid=8355;			
NP	[1]			
NP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Embryo;			
RX	MEBLIN=22388257; PubMed=12477932; DOI=10.1073/pnas.2426038999;			
RX	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,			
RA	Alausner S.F., Zeng B., Buettow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Ueda T.B., Toshynski S., Carrinetti P., Prange C.,			
RA	Rata S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,			
RA	Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,			
RA	Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Faney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Boulford G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimmwood J., Schmutz J., Myers R.M.,			
RA	Butterfield A.Y.S.N., Krzyzinski M.I., Skalska U., Smallie D.E.,			
RA	Schnecker F., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Embryo;			
RX	MEBLIN=22381132; PubMed=12454917; DOI=10.1002/dvdy.10174;			
RA	Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,			
RA	Richardson P.,			
RT	"Genomic and genomic tools for Xenopus research: The NIH Xenopus			
RT	Initiative."			
RL	Dev. Dyn. 225:384-391(2002).			
RN	[3]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Embryo;			
RL	Klein S., Strausberg R.;			
RL	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; BC068805; AA068805.1; -, mRNA.			
DR	GO; GO:0005524; F:ATP binding; IEA.			
DR	GO; GO:0004574; P:protein serine/threonine kinase activity; IEA.			
DR	GO; GO:0004713; P:protein-tyrosine kinase activity; IEA.			
DR	GO; GO:0016740; P:transferase activity; IEA.			
DR	GO; GO:0051301; P:cell division; IEA.			
DR	GO; GO:0006468; P:protein amino acid phosphorylation; IEA.			
DR	InterPro; IPR000719; Prot_kinase.			
DR	InterPro; IPR002230; Ser_thr_kinase.			
DR	InterPro; IPR001245; Tyr_kinase.			
DR	Pfam; PF00069; Pkinase; 1.			
DR	ProDom; PD000001; Prot_kinase; 1.			
DR	SMART; SM00220; S_TKc; 1.			
DR	SMART; SM00219; Tyrc; 1.			
DR	PROSITE; PSS0011; PROTEIN_KINSE_DOM; 1.			
DR	SEQUENCE 526 AA; 5936 MW; E966FC5FCBC742B CRC64;			

Query Match	56.7%;	Score 1514;	DB 2;	Length 526;
Best Local Similarity	58.8%;	Pred. No. 8.1e-101;		

Matches 303; Conservative 68; Mismatches 108; Indels 36; Gaps 7;

QY 1 MAAPAPARAREREREREDSESDILESPPCGRNQKREYNQNMPLQSTFLAMD 60
 Db 28 VSNP-RP-----PQEEBESBESBESILESPPCGRNQKREYNQNMPLQSTFLAMD 81
 QY 61 TEBGVEVWNNELHFGDKAPAAHEKIQVFEQVLVDHNPVIVLHKYMLDTSACARVI 120
 Db 82 TEBGVEVWNNELHFGDKAPAAHEKIQVFEQVLVDHNPVIVLHKYMLDTSACARVI 141
 QY 121 FITEYVSSGSKQFLKTKTKNKKANARAWKRWCTQILSALSFHACSPPIHGNLTSDT 180
 Db 142 FITEYVSSGSKQFLKTKTKNKKANARAWKRWCTQILSALSFHACSPPIHGNLTSDT 201
 QY 181 IFIOHNLKIGSVYMHIFSNALRPPTALPDLSPIRAERERLNHFFPEYGEVAD- 239
 Db 202 IFIOHNLKIGSVY-----ADPTINNHYKTCRECKSLHFFAPEYGEVATNV 247
 QY 240 GTAVDIFSPGCMALMAVLEIQTINGDTR-VTEAIAARSHLSDPNKREFILCCLADPA 298
 Db 248 TTVNDIYSPGCMALMAVLEIQNGESSYVQEAINNAIQFLBPLQREFIQKLETDPS 307
 QY 299 RPSAHSLEPHRYLFEVHSLKLLAHCFIGHQYLMPEVNVBEKTKAMDIAVAELRPR 358
 Db 308 KRPTARELLFHQALFEVPSLKLAAHCIVGHQMIANALEITKNLDMASAVLAETHTD 367
 QY 359 RPLQMRYSVSEVSEFMEIDKFLIEDVANGIYPLMNFATPPLGLPRLAPPPEVQKA----- 413
 Db 368 RQVGRMLFSQSPALBIDKFLIEDVANGIYPLTAFGVPLP-----QQPQOEYVKSPPVPP 420
 QY 414 --KTPEPDPSESTRKYIQMCNLERSEDKARWMLTLLVLVEDRLHQLTYYDLPTDQAQ 471
 Db 421 SVTTPPEPAVEVERKRVQVQMCNIESYDEGAKHHTLLTKLEDKLNHLSCDLLPNNIQ 480
 QY 472 DLASELVHYGFLHEDDMKLAFLSTFLKYRGTO 506
 Db 481 ELAAELVQLGFSVSECDQSRILTCLEDAFSKPYFTR 515

RESULT 12
 ID Q5SNQ3_BRARE PRELIMINARY; PRT; 519 AA.
 AC Q5SNQ3;
 DT 01-FEB-2005 (T-EMBLrel. 29, Created)
 DT 01-FEB-2005 (T-EMBLrel. 29, Last sequence update)
 DE Novel protein similar to vertebrate nuclear receptor binding protein (NRBP).
 GN ORFNames=DKEY-12H9, 7-001;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 NC NCB1_Taxid=7955;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Phillimore B.;
 RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL954831; CA120617.1; -; Genomic_DNA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0004682; F:receptor activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR Interpro; IPR002290; Ser Thr kinase.
 DR Interpro; IPR001245; Tyr_kinase.
 DR Pfam; PF00069; Kinase; 1.
 DR Prodom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 KM Receptor.

SQ SEQUENCE 519 AA; 58644 MW; CDPD99609B0C18C4 CRC64;

Query Match 55.7%; Score 1486.5; DB 2; Length 519;
 Best Local Similarity 58.7%; Pred. No. 7.7e-99;

Matches 297; Conservative 66; Mismatches 106; Indels 37; Gaps 7;

QY 5 EPAPRRAREREREREDSESDILESPPCGRNQKREYNQNMPLQSTFLAMDTEEG 64
 Db 29 QPAP-----SATDEBESBESBESILESPPCGRNQKREYNQNMPLQSTFLAMDTEEG 84
 QY 65 VEYVWNNELHFGDKAPAAHEKIQVFEQVLVDHNPVIVLHKYMLDTSACARVITE 124
 Db 85 VEYVWNNELHFGDKAPAAHEKIQVFEQVLVDHNPVIVLHKYMLDTSACARVITE 144
 QY 125 YVSSGSKQFLKTKTKNKKANARAWKRWCTQILSALSFHACSPPIHGNLTSDTIFQ 184
 Db 145 YVSSGSKQFLKTKTKNKKANARAWKRWCTQILSALSFHACSPPIHGNLTSDTIFQ 204
 QY 185 HNGLIKIGSVYMHIFSNALRPPTALPDLSPIRAERERLNHFFPEYGEVAD-GTAV 243
 Db 205 HNGLIKIGSVY-----ADPTINNHYKTCRECKSLHFFAPEYGEVATNV 250
 QY 244 DIFSPGCMALMAVLEIQTINGDTR-VTEAIAARSHLSDPNKREFILCCLADPARPS 302
 Db 251 DIFSPGCMALMAVLEIQNGESSYVQEAINNAIQFLBPLQREFIQKLETDPSKRP 310
 QY 303 AHSLEPHRYLFEVHSLKLLAHCFIGHQYLMPEVNVBEKTKAMDIAVAELRPRRPL 362
 Db 311 ARELLFHQALFEVPSLKLAAHCIVGHQMIANALEITKNLDMASAVLAETHTD 365
 QY 363 QMRYSVSEVSEFMEIDKFLIEDVANGIYPLMNFATPPLGLPRLAPPPEVQ-----KAKTP 416
 Db 366 QKLQSPALBIDKFLIEDVANGIYPLTAF-----GMPCCQOQOEYVKSPPVPSKTP 419
 QY 417 TPEPDPSESTRKYIQMCNLERSEDKARWMLTLLVLVEDRLHQLTYYDLPTDQAQ 476
 Db 420 TPEPDPSESTRKYIQMCNIESYDEGAKHHTLLTKLEDKLNHLSCDLLPNNIQ 479
 QY 477 LVHYGFLHEDDMKLAFLSTFLKY 502
 Db 480 LVQLGFSVSECDQSRILTCLEDAFSKPYFTR 505

RESULT 13
 ID Q4S1G5_TETNG PRELIMINARY; PRT; 501 AA.
 AC Q4S1G5;
 DT 13-SEP-2005 (T-EMBLrel. 31, Created)
 DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
 DE Chromosome 6 SCAF14768, whole genome shotgun sequence.
 GN ORFNames=GSTENG0025570001;
 OS Tetradon nigrovittatus (Green puffer).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Tetraodontidae; Tetraodon.
 NC NCB1_Taxid=99883;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Uallion O., Aury J.W., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Maucel E., Bouneau L., Fischer C., Ozout-Coetzee C., Bernot A.,
 RA Nicard S., Jaffe D., Fisher S., Iulfalga G., Dosset A., Segurens B.,
 RA Desliva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
 RA Anthouard V., Jubin C., Castell V., Katinka M., Vacherie B.,
 RA Biemont C., Skalli Z., Cattolico L., Poulin J., De Berardinis V.,
 RA Cnaud C., Duprat S., Broclet P., Coutanceau J.P., Gouzy J.,
 RA Parra G., Lardier G., Chappie C., McKernan K.J., McSwan P., Bosak S.,
 RA Kellis M., Volff J.N., Guigo R., Zody M.C., Mesitov J.,
 RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Lander V., Schacher V., Queller F., Saurin W., Scarpelli C.,
 RA Wincker P., Lander E.S., Weissenbach J., Roest Crollins H.;

RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype." ;
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -1- FUNCTION: Plays a key role in the control of the eukaryotic cell
CC cycle. It is required in higher cells for entry into S-phase and
CC mitosis. Component of the kinase complex that phosphorylates the
CC repetitive C-terminus of RNA polymerase II. Catalytic component of
CC MPF (By similarity).
CC -1- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in
CC mature oocytes (By similarity).
CC EMBL; CAE01014768; CAG0517.1; -; Genomic_DNA.
CC InterPro: IPR000719; Prot_Kinase.
CC InterPro: IPR002290; Ser_Thr_kinase.
CC InterPro: IPR001245; Tyr_kinase.
CC Pfam: PF00669; Pkinase; 1.
CC ProDom: PD000001; Prot_Kinase; 1.
CC SMART: SM00220; S_TKc; 1.
CC SMART: SM00219; TYKc; 1.
CC PROSITE: PSS0011; PROTEIN_KINASE_DOM; 1.
FT NON_TER 1 501
FT 501 501
SQ SEQUENCE 501 AA; 57651 MW; C28F26487BA61567 CRC64;

Query Match 52.5%; Score 1400.5; DB 2; Length 501;
Best Local Similarity 55.3%; Pred. No. 1.2e-92;
Matches 280; Conservative 55; Mismatches 83; Indels 87; Gaps 5;

QY 43 QVQNGMNGIGQSTFLAMDTEEGVGVVNNELHPSGRKAFANBEKIQTEFQVLVVDHPNI 102
DB 1 QVQNGVNPVGEVASILAMDTEEGVGVVNNELVFSDKVFKFQKEBKIKEMFENLQVHEHPNI 60
QY 103 VKLHKWLDTSEACARVIFTEYVSSGSLKQFLKTKKHKANNAKMKWCQIIISAL- 161
DB 61 VKFHKWLDMKESQARVIFTEYVSSGSLKQFLKTKKHKNTNVAKMKWCQIIISALR 120
QY 162 -----SFLHACSPRIHGNLTSDTIFIOHNLKIGSYMR 197
DB 121 YSILYIMRFLMTSPDDLFPQSYIHSQDPRIHGNLTCTITFIOHNLKIGSYMR 180
QY 198 IFSNALRPTALPD-DLRSEIRAREELRNLFPPRYGEVA--DGTAVDIFSGMCALE 254
DB 181 LFVN-----VFPDASVHGKGRORHDEQRNLHFFAPRYGKSGEDVYADIFSGICALE 234
QY 255 MATLETQNGDTVTVEALIRARHSLSDPNMR----- 286
DB 235 MAVLETIQANGDSVYSKALVNAQHSLEDPMRYSVNRTPCPRYTSQOGMLYITLLTRL 294
QY 287 -----EFILCCLARPARPSASLLFHRVLFVHSLKLL 321
DB 295 CSLISCSWNPETEGCNFTLDFLOEFTQSCLRHNAKLRPTADLLFHRVLFVHSLKLL 354
QY 322 AAHCFIQHOYLMBENVVEEKTAMDIAVLAELPRPRRPLOWRYSVEVSFMEIDKLEVDY 381
DB 355 AAHCLINNOYLPLENCEEEKTKSIDPAVVAELIRHNDROGVQAKYHVSFLEIDKLEVDY 414
QY 382 RNCIYPLANFAPRPLGLPRVLAAPRPEVOKATPTPEPDSSTRKYIOMCNLERSESDK 441
DB 415 KNGIYPLANFASMPHPRPALSLSGOVETVKTPTPEPETETRKVQVQHCNLESNEEG 474
QY 442 ARWHLTLVLIEDRLRQLTYDILLPT 467
DB 475 TKTHLSLFLKMDDKLHNLSCDIFPS 500

RESULT 14
Q9NSY0_HUMAN

ID Q9NSY0_HUMAN PRELIMINARY; PRT; 258 AA.
AC Q9NSY0;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-FEB-2005 (T-EMBLrel. 29, Last annotation update)
DE Hypothetical protein DKFZP434P086 (LOC340371 protein).
GN Name=DKFZP434P086; Synonyms=LOC340371;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Testis;
RA Koehler K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
RE EMBL; AL137662; CAB70864.1; -; mRNA.
DR EMBL; CR457350; CAG33631.1; -; mRNA.
DR PIR; T46491; T46491.
DR GO; GO:0005524; F-ATP binding; IEA.
DR GO; GO:0004672; F-protein kinase activity; IEA.
DR GO; GO:0006468; F-protein amino acid phosphorylation; IEA.
DR InterPro: IPR000719; Prot_Kinase.
DR ProDom: PD000001; Prot_Kinase; 1.
KW Hypothetical protein; Nucleotide-binding.
SQ SEQUENCE 258 AA; 29852 MW; 3C886AAD4CDE26EA CRC64;

Query Match 50.4%; Score 1346; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 4.4e-89;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 MCALEAVALEIQNGDTRVTEAIAARHSLSDPNREFTLCCIARDPARPSAHSLLPH 309
DB 1 MCALEAVALEIQNGDTRVTEAIAARHSLSDPNREFTLCCIARDPARPSAHSLLPH 60
QY 310 RVLFEVHSLKLLAAHCFIQHOYLMBENVVEEKTAMDIAVLAELPRPRRPLOWRYSVEY 369
DB 61 RVLFEVHSLKLLAAHCFIQHOYLMBENVVEEKTAMDIAVLAELPRPRRPLOWRYSVEY 120
QY 370 SFMEIDKLEVDYRNGIYPLANFAPRPLGLPRVLAAPRPEVOKATPTPEPDSSTRKYI 429
DB 121 SFMEIDKLEVDYRNGIYPLANFAPRPLGLPRVLAAPRPEVOKATPTPEPDSSTRKYI 180
QY 430 OMOCNLERSESDARWHLTLVLIEDRLRQLTYDILLPTQSADLASLVHGYGLHEDDM 489
DB 181 OMOCNLERSESDARWHLTLVLIEDRLRQLTYDILLPTQSADLASLVHGYGLHEDDM 240
QY 490 KLAAPLESTFLKYRGTOA 507
DB 241 KLAAPLESTFLKYRGTOA 258

RESULT 15
Q9ARKP5_TETNG
ID Q9ARKP5_TETNG PRELIMINARY; PRT; 483 AA.
AC Q9ARKP5;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
DE Chromosome 5 SCAR15026, whole genome shotgun sequence.
DE (Fragment).
GN ORFname=GSTENG00032841001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;

RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Jellison O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
 RA Mucelli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
 RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
 RA Desliya C., Salenoubat M., Levy M., Boudet N., Castellano S.,
 RA Authouard V., Jubin C., Castell V., Katinka M., Vacherie B.,
 RA Blemont C., Skalli Z., Cattolico L., Poulsen J., De Bernardis V.,
 RA Cuand C., Duprat S., Brottier P., Couanceau J.P., Gouy J.,
 RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
 RA Kallis M., Volt JN., Guigo R., Zody M.C., Mesirov J.,
 RA Landrad-Ton K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
 RA Landrad-Ton K., Schachter V., Queller F., Saurin W., Scarpelli C.,
 RA Winkler P., Lander E.S., Weisenbach J., Roest Crolius H.,
 RT "genome duplication in the teleost fish Tetraodon nigroviridis reveals
 the early vertebrate proto-karyotype.",
 RL Nature 431:946-957(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RG Genoscope: Whitehead Institute Centre for Genome Research;
 RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 preliminary data.
 CC -!- FUNCTION: Plays a key role in the control of the eukaryotic cell
 cycle. It is required in higher cells for entry into S-phase and
 mitosis. Component of the kinase complex that phosphorylates the
 repetitive C-terminus of RNA polymerase II. Catalytic component of
 MPF (By similarity).
 CC -!- SUBUNIT: Forms a stable but non-covalent complex with cyclin B in
 mature oocytes (By similarity).
 CC EMBL: CAAB01015026; CAG11037.1; -, Genomic_DNA.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; Kinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TYKc; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 FT NON_TER 1
 FT NON_TER 483
 SQ SEQUENCE 483 AA; 54608 MW; 1A970E4D3EFA4224 CRC64;
 Query Match 49.9%; Score 1331.5; DB 2; Length 483;
 Best Local Similarity 53.6%; Pred. No. 1,1e-87;
 Matches 265; Conservative 70; Mismatches 114; Indels 45; Gaps 4;
 QY 19 EDESEDESDILLESPPCGRMQRQVQGMPPGLQSTFLAMPDEEGVVMNELLHPGDRK 78
 DB 8 DEESDDESEILLESPPCGRMQRQVQGMPPGLQSTFLAMPDEEGVVMNELLHPGDRK 67
 QY 79 AFAAHEKIQIVPEQVLVDHPNIVKLKTYWLDTSEACARVIFTEYVSSGSLKQPLKKT 138
 DB 68 NFKQLEKVKAVPDNLHLHANIYKFKYKWDTKGRARAVIFTEYVSSGSLKQPLKKT 127
 QY 133 KKHAKAMNAAKMKWCTQIIISAL-----SFLHACSPPII 172
 DB 128 KKHAKMNERKALKMKWCTQIIISALKTLSVQALGDVAKSSILMFSPICSYLHSDPPII 187
 QY 173 HGNLTSPDTFIQHNGLIKISVMHRIFSNMLRPPALPDDLRSPIRAAREELNTHFPF 232
 DB 188 HGNLTCDTIFIQHNGLIKISV-----APDTINNHVKTCYEEQKVLHFPYAP 233
 QY 233 EYGEVADGTAVDIFSPGMCALEMAVLEIQTNGDTR-VTEBAIARASHLSDPNMRPFIIC 291
 DB 234 EYGDNDTAVDIFSPGMCVLEMAVLEIHONGSSSYSDAINNAILDLPLQKELIQK 293
 QY 292 CLARDPARPSANSLFHRVLFYVHSLKLAACFIQHOYILMPENVVEEKTAMDLHAVL 351
 DB 294 CLESDEPVRPTARELLFDPALFEVPLKLLAASIVHQYIMENALEEMTKNLDPNLVI 353
 QY 352 AELPRPRRPPLOMRYSVSEVMELDKFLIEDVRNGIYPLMNPATRPGLPRVLAPRPEVQ 411

DB 354 AEM-----KEDIQKLSQFPALIEDKFLIEDVRNGIYPTATGCLSPQKPOQGEVVKSPVVR 409
 QY 412 KAKTPPEPDSSTRKYIQMCNLERSEDKARWHTLLVLVEDRLHRLTYDILPTDSAQ 471
 DB 410 LVSPPTPEPALLETFRVQMCNLEIEEGTKFHTLLKLEDRDLNHLSCDMLPNESVQ 469
 QY 472 DLASELVHGFLEH 485
 DB 470 ELAGELVLELISE 483

Search completed: January 12, 2006, 11:21:58
 Job time : 78 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2006, 11:22:06 ; Search time 115 Seconds
(without alignments)
1842.082 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670
Sequence: 1 MAPEPAPRRARERERED.....RMKLAFLSTFLKNGTQA 507

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Published Applications AA Main:
1: /cgn2_6/ptodaca/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodaca/1/pubppaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodaca/1/pubppaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodaca/1/pubppaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodaca/1/pubppaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodaca/1/pubppaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2670	100.0	507	4	US-10-618-941-99
2	2499	93.6	499	5	US-10-840-512-185
3	1580	59.2	326	3	US-09-862-027-10
4	1580	59.2	326	3	US-10-989-228-10
5	1534.5	57.5	335	3	US-09-840-787-20
6	1534.5	57.5	335	4	US-10-153-668-322
7	1534.5	57.5	335	4	US-09-764-875-749
8	1529.5	57.3	335	4	US-10-059-585-2
9	1508.5	56.5	535	4	US-10-040-647-10
10	1444	54.1	501	3	US-09-925-297-609
11	1310	49.1	434	4	US-10-153-668-434
12	1174.5	44.0	637	6	US-11-097-143-270
13	946	35.4	340	3	US-09-764-868-798
14	946	35.4	340	3	US-09-764-875-1053
15	882.5	33.1	752	4	US-10-292-798-1714
16	882.5	33.1	886	4	US-10-017-161-2068
17	756	28.3	302	3	US-09-764-868-822
18	756	28.3	302	3	US-09-764-875-750
19	543	20.3	231	3	US-09-862-027-36
20	543	20.3	231	5	US-09-862-027-35
21	415.5	15.6	461	3	US-09-862-027-35
22	415.5	15.6	461	5	US-09-862-027-35
23	395	14.8	779	4	US-10-353-529-45
24	395	14.8	1345	4	US-10-433-794-17
25	395	14.8	2193	5	US-10-490-592-5
26	395	14.8	2245	4	US-10-618-941-108
27	392	14.7	2132	5	US-10-840-512-195

28	386	14.5	1751	5	US-10-840-512-197	Sequence 197, App
29	378	14.2	1069	4	US-10-182-243-50	Sequence 50, App1
30	378	14.2	1743	4	US-10-460-545-24	Sequence 2, App1
31	376.5	14.1	1234	4	US-10-052-648A-36	Sequence 36, App1
32	372	13.9	516	3	US-09-862-027-34	Sequence 34, App1
33	372	13.9	516	5	US-10-989-228-34	Sequence 34, App1
34	371.5	13.9	1243	4	US-10-478-146A-37	Sequence 8, App1
35	370.5	13.9	1231	4	US-10-052-648A-37	Sequence 37, App1
36	370.5	13.9	1243	4	US-10-196-935A-4	Sequence 4, App1
37	368.5	13.8	1601	3	US-09-862-027-40	Sequence 40, App1
38	368.5	13.8	1601	3	US-10-989-228-40	Sequence 40, App1
39	367.5	13.8	557	4	US-10-052-648A-12	Sequence 12, App1
40	367.5	13.8	1251	4	US-10-114-270-80	Sequence 80, App1
41	365.5	13.7	658	4	US-10-424-599-170860	Sequence 170860, App1
42	365.5	13.7	746	4	US-10-425-114-72414	Sequence 72414, A
43	363.5	13.6	601	4	US-10-437-963-103929	Sequence 103929, App1
44	359.5	13.5	670	4	US-10-052-648A-38	Sequence 38, App1
45	359.5	13.5	2108	5	US-10-491-467-2	Sequence 2, App1

ALIGNMENTS

RESULT 1

US-10-618-941-99

Sequence 99, Application US/10618941

Publication No. US2004019792A1

GENERAL INFORMATION:

APPLICANT: MANNING, GERARD

APPLICANT: CAENEPEL, SEAN

TITLE OF INVENTION: NOVEL KINASES

FILE REFERENCE: 034536-0321

CURRENT APPLICATION NUMBER: US/10/618,941

PRIOR APPLICATION DATE: 2003-07-15

PRIOR FILING DATE: 2002-07-15

NUMBER OF SEQ ID NOS: 143

SOFTWARE: PatentIn version 3.2

SEQ ID NO 99

LENGTH: 507

TYPE: PRT

ORGANISM: Homo sapiens

US-10-618-941-99

Query Match 100.0%; Score 2670; DB 4; Length 507;
Best Local Similarity 100.0%; Pred. No. 1.4e-220;
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAPEPAPRRAREREREREDSESDIIEBPCGRWQRRQVNGNMPGLQSTFLAND 60
DB	1	MAPEPAPRRAREREREREDSESDIIEBPCGRWQRRQVNGNMPGLQSTFLAND 60
QY	61	TEBGEVNVNNEHFGGRKAPFAHEEKTQVFQOLVVDHPNIVKHLKWLDTSEACARYI 120
DB	61	TEBGEVNVNNEHFGGRKAPFAHEEKTQVFQOLVVDHPNIVKHLKWLDTSEACARYI 120
QY	121	FTTEVYSSGSLKQFLKTKKNKAMARAKKWCQIILSALFSLHACSPPIHGNLTSDT 180
DB	121	FTTEVYSSGSLKQFLKTKKNKAMARAKKWCQIILSALFSLHACSPPIHGNLTSDT 180
QY	181	IFIQHNGLIKIGSVHRIFSNARIPPTALPDDLRSPIRAERELRNHFPPEYGEVADG 240
DB	181	IFIQHNGLIKIGSVHRIFSNARIPPTALPDDLRSPIRAERELRNHFPPEYGEVADG 240
QY	241	TAVDIFSPGMCALMAVAVIEIQNGDRTVTEALIAARHSLSPNNREFITLCLAPRAR 300
DB	241	TAVDIFSPGMCALMAVAVIEIQNGDRTVTEALIAARHSLSPNNREFITLCLAPRAR 300
QY	301	PSAHSILFRRVLFEVHSLKILAAHCFIQOYIMPEVVEKTKAMDIAVLAELPRPRRP 360
DB	301	PSAHSILFRRVLFEVHSLKILAAHCFIQOYIMPEVVEKTKAMDIAVLAELPRPRRP 360

Qy	361	PLQMRYSVEVFMELDFLEJVEVRNGIYPLANNPATPELGIPLRAPPEVEOKATTPPER	420
Db	361	PLQMRYSVEVFMELDFLEJVEVRNGIYPLANNPATPELGIPLRAPPEVEOKATTPPER	420
Qy	421	FDSTFRVYIOMQCNLRSRSEDKARWHITLLVYLEDRLHQLTYDILPTMSADLASLTVHY	480
Db	421	FDSTFRVYIOMQCNLRSRSEDKARWHITLLVYLEDRLHQLTYDILPTMSADLASLTVHY	480
Qy	481	GFLEHEDDRMKLAFLFLESTFLKRGTOA	507
Db	481	GFLEHEDDRMKLAFLFLESTFLKRGTOA	507

```

US-10-840-512-185
RESULT 2
Sequence 185, Application US/10840512
Publication No. US20050125852A1
GENERAL INFORMATION:
APPLICANT: CAENEDEEL, SEAN
APPLICANT: MANNING, GERRARD
APPLICANT: CHARVDCZAK, GLEN
APPLICANT: GRIGORIEV, IGOR
TITLE OF INVENTION: NOVEL KINASES
FILE REFERENCE: 034536-1455
CURRENT APPLICATION NUMBER: US/10/840,512
CURRENT FILING DATE: 2004-05-07
PRIOR APPLICATION NUMBER: 60/469,014
PRIOR FILING DATE: 2003-05-09
NUMBER OF SEQ ID NOS: 239
SOFTWARE: PatentIn version 3.2
SEQ ID NO 185
LENGTH: 499
TYPE: prt
ORGANISM: Mus musculus
US-10-840-512-185

```

Query Match	93.6%	Score 2499	DB 5	Length 499
Best Local Similarity	93.9%	Pred. No. 6.8e-206		
Matches	17	Mismatches	12	Indels 8
				Gaps 2

Oy	MAAPEPRARAREERREDESEDESDILIEBSPCCRMQKREQVNOGMPLOSTFLAMD	60
Db	1 MAAPERPARG--REBEREDSEDESDILIEBSPCCRMQKREQVNOGMPLOSTFLAMD	58
Oy	TEBGEVYVNMELHFQGRKAPAAHEEKIQTFVEQVLVDHPNIVKLHKWLDTSEACARVI	120
Db	TEEBGEVYVNMELHFQGRKAPAAHEEKIQTFVEQVLVDHPNIVKLHKWLDASEARAVI	118
Oy	PTEYVSSGSLKQFLKTKTKKHKAMARAKMKWCQIILSLSPLHACSPPIIHENLISDT	180
Db	119 PTEYVSSGSLKQFLKTKTKKHKAMARAKMKWCQIILSLSPLHACSPPIIHENLISDT	178
Oy	181 IFIQHNLKIGISVMHRIFSNALRPETALPDDLRSPIABREBELRNLIHFPEEYGEVADG	240
Db	179 IFIQHNLKIGISVMYRIFSN-----ALPDLRSPIABREBELRNLIHFPEEYGEVNDG	232
Oy	241 TAYDIBSEFGKCALEMVALEIQTNQDTRVTEEAIAARHSLSDPNNKREIILCLLARDARR	300
Db	233 TAYDIBSEFGKCALEMVALEIQAGDTRVTEEAIAARHSLSDPNNKREIILCLLARDARR	292
Oy	301 PSAHSLLFHRLVEVHSLKLAAHCFIQHQYLPMPENVVEEYTKAMDILHAYLAELPRPRP	360
Db	293 PSAHNLFLHRLVEVHSLKLAAHCFIQHQYLPMPENVVEEYTKAMDILHAYLAEMRQHPG	352
Oy	361 PLQWRYSVEFMELOKFLBEDVRNGIYPLMNFATPTPLGPRYLARPPEEYOKATPTPEP	420
Db	353 PMQWRYSVEFLELDKFLBEDVRNGIYPLMNFFAAAPTPLGPRYLARPPEEAKATPTPEP	412
Oy	421 FDESTRVIVIMOCNLRSESDKARWHLTLLVVEDRLHQLRYDILLPTDQSAODLASLELVHY	480
Db	413 FDESTRVIVIMOCNLRSESDKARWHLTLLVVEDRLHQLRYDILLPTDQSAODLAELVHY	472
Oy	481 GFLHEDRMKLAFLSTFLKTRGTQAA	507

Db 473 GFLHEDDRTKLAFLFTTLKRYGTQA 499

```

RESULT 3
US-09-862-027-10
Sequence 10, Application US/09862027
Patent No. US20020142428A1
GENERAL INFORMATION:
APPLICANT: Hodge, Martin R.
TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
FILE REFERENCE: 35800/234862
CURRENT APPLICATION NUMBER: US/09/862,027
CURRENT FILING DATE: 2001-05-21
PRIOR APPLICATION NUMBER: US 09/745,473
PRIOR FILING DATE: 1999-06-30
NUMBER OF SEQ. ID NOS: 82
SOFTWARE: FastSeq For Windows Version 4.0
SEQ ID NO 10
LENGTH: 326
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(326)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-862-027-10

```

Query Match	59.2%	Score 1580	DB 3	Length 326
Best Local Similarity	97.4%	Pred. No. 4e-127		
Matches 301	0	Mismatches 2	Indels 6	Gaps 1

```

QY      1  MAAPPAAPRAAREREREREDSESDSILLES PCGRQOKRREQVNOGSMGLQSTFLAMD 60
Db      12  MAAPPAAPRAAREREREREDSESDSILLES PCGRQOKRREQVNOGSMGLQSTFLAMD 71
QY      61  TEEGVEVWNNEHLFGDRKAFAPAAHEEKIQTVEQVLVDHPNIVKLAKYMLDTSBACARVI 120
Db      72  TEEGVEVWNNEHLFGDRKAFAPAAHEEKIQTVEQVLVDHPNIVKLAKYMLDTSBACARVI 131
QY      121  FITEEYVSSSSLQPLKTKTKKHKKANNAAMKRWCTQILSALSFHACSPPIHGNLTSDT 180
Db      132  FITEEYVSSSSLQPLKTKTKKHKKANNAAMKRWCTQILSALSFHACSPPIHGNLTSDT 191
QY      181  IFIOHNGLIKIGISVWHRIFSNALRBPPTALPDDLSPITRAEREBLRNLFPPPEXGEVADG 240
Db      192  IFIOHNGLIKIGISVWHRIFSN-----ALPDDLSPITRAEREBLRNLFPPPEXGEVADG 245
QY      241  TAVDIFSPGMCALEMAVLEIQTNGDTRVTEBIAIARARHSISDPNMRSEFIIICCLARDPARR 300
Db      246  TAVDIFSPGMCALEMAVLEIQTNGDTRVTEBIAIARARHSISDPNMRSEFIIICCLARDPARR 305
QY      301  PSASHSLFF 309
Db      306  PSVHSLFF 314

RESULT 4
US-10-989-228-10
/ Sequence 10, Application US/10989228
/ Publication No. US20050089917A1
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin R.
/ TITLE OF INVENTION: Novel Kinases and Uses Thereof
/ FILE REFERENCE: 35800/234862
/ CURRENT APPLICATION NUMBER: US/10/989, 228
/ CURRENT FILING DATE: 2004-11-15
/ PRIOR APPLICATION NUMBER: US/09/862, 027
/ PRIOR FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 09/345, 473
/ PRIOR FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 82
/ SOFTWARE: FastSeq for Windows Version 4.0

```



```

/ PRIOR APPLICATION NUMBER: PCT/JP00/05060
/ PRIOR FILING DATE: 2000-07-28
/ PRIOR APPLICATION NUMBER: US 60/183,322
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: US 60/159,590
/ PRIOR FILING DATE: 1999-10-18
/ PRIOR APPLICATION NUMBER: JP 2000-118776
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: JP 2000-183767
/ PRIOR FILING DATE: 2000-05-02
/ PRIOR APPLICATION NUMBER: JP 11-248036
/ PRIOR FILING DATE: 1999-07-29
/ NUMBER OF SEQ ID NOS: 64
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 535
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-059-585-2

```

Query Match	57.3%	Score 1529.5;	DB 4;	Length 535;
Best Local Similarity	61.4%;	Pred. NO. 1.7e-122;		
Matches 304; Conservative	59;	Mismatches 103;	Indels 29;	Gaps 6;

Dd	44	EEBESBDSSEJLEBSPCGRMQGRREVQVGRVPGIDASYLAMDTEBEVEVMNNEVOFSE	103
Qy	17	ERDSEDSDESDILEBSPCGRMQGRREVQVQGNMPLGISTFLAMDTREEVEVVMNNEHJGD	76
Qy	17	ERDSEDSDESDILEBSPCGRMQGRREVQVQGNMPLGISTFLAMDTREEVEVVMNNEHJGD	76
Dd	44	EEBESBDSSEJLEBSPCGRMQGRREVQVGRVPGIDASYLAMDTEBEVEVMNNEVOFSE	103
Qy	77	RKAFAAHEEKIQVFEQVLVDHPNI VKLAKTWTDSCEARVIFTEYVSSGSLKOPFK	136
Dd	104	RKYKJQOEELVVRVFNDLQLBHLNVKFKYMWADIKENKARVIFTEYVSSGSLKOPFK	163
Qy	137	KTKGNKAMAMARAKMKTCTQILSALSFLACSPPIHGNLTSDTPIQHNGLKIGSVMH	196
Dd	164	KTKGNKHTNEMKMKMKTCTQILSALSFLACSPPIHGNLTCDTPIQHNGLKIGSV--	221
Qy	197	RIFSNA LRPTLPDLRSPIRAEREELNLHFPPPEYGEVAD- GTAVDIFSFGKALEM	255
Dd	222	-----APDITNNHVKTCEBOQNHFHFPPEYGEVYNNVTAVDIFSFGKALEM	268
Qy	256	AVLEIQNGDTR-VTEBAIARARHSISDPNMBEFLCCIARDPARRPSAHSILPHRVLFE	314
Dd	270	AVLEIQNGESSVVRPEALISSAQLLEDBLQNEBFIQKLOSEPARRPARRELNLHPALFE	328
Qy	315	VHSLSKLLAAHCFIQHOYLPMBENVVEBKTAMDHLAVLAELPR-PRRPELOMRYSVSPME	373
Dd	330	VPSLKILLAHCHICGHOMIPENALBEITKMDTSAVLAEIPAGFORBEVQTLYSQSPALE	389
Qy	374	LDBFELDVANGIYELMNFPAATRPLGIPRYLAPRPEV-----QKAKTPPEPDSSTRK	427
Dd	390	LDBFELDVANGIYELPLAF-----GLPRQOQOEBSVSPVVPBSVKPIPEERAEVETRK	443
Qy	428	VVICQGNLERSBDKAMWLTLLVLVEDLRHLQTLVDLPTDSADLASELVHYGFTLEDD	487
Dd	444	VVICQGNISVESVEGVGHITLLLKLEDKSRHLSCDLMPNBNIPELAAELVQGFISEAD	503
Qy	488	RMLTAAFLSESTPLKY	502
Dd	504	QSRULTSLEETLANKF	518

RESULT 9
US-10-040-647-10
/ Sequence 10, Application US/10040647
/ Publication No. US20030092154A1
/ GENERAL INFORMATION:
/ APPLICANT: (US only) ANTALIS Toml Marie and HOOBER John David
/ TITLE OF INVENTION: NOVEL MOLECULES
/ NUMBER OF SEQUENCES: 30
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
/ STREET: 400 GARDEN CITY PLAZA
/ CITY: GARDEN CITY

STATE: NEW YORK
COUNTRY: USA
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/040,647
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/023,942
FILING DATE: <Unknown>
APPLICATION NUMBER: P05101/97
FILING DATE: 13-FEB-1997
APPLICATION NUMBER: PP0422/97
FILING DATE: 18-NOV-1997

APPLICATION NUMBER: International PCT Application
FILING DATE: 13-FEB-1998

ATTORNEY/AGENT INFORMATION:
NAME: DIGILIO, FRANK S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 11168

TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742 4343
TELEFAX: (516) 742 4366
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 535 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-10-040-647-10

Query Match	56.5%	Score 1508.5	DB 4	Length 535
Best Local Similarity	60.8%	Pred. NO. 1.1e-120		
Matches 301; Conservative	60;	Mismatches 105;	Indels 29;	Gaps 6;

[illegible]


```
/ GENERAL INFORMATION:
/ APPLICANT: Venter, J. Craig
/ TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
/ TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
/ TITLE OF INVENTION: DROSOPHILA GENES.
/ FILE REFERENCE: CL000728
/ CURRENT APPLICATION NUMBER: US/11/097,143
/ CURRENT FILING DATE: 2005-04-04
/ PRIOR APPLICATION NUMBER: 60/157,832
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: 60/160,191
/ PRIOR FILING DATE: 1999-10-19
/ PRIOR APPLICATION NUMBER: 60/161,932
/ PRIOR FILING DATE: 1999-10-28
/ PRIOR APPLICATION NUMBER: 60/164,769
/ PRIOR FILING DATE: 1999-11-12
/ PRIOR APPLICATION NUMBER: 60/173,383
/ PRIOR FILING DATE: 1999-12-28
/ PRIOR APPLICATION NUMBER: 60/175,693
/ PRIOR FILING DATE: 2000-01-12
/ PRIOR APPLICATION NUMBER: 60/184,831
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: 60/191,637
/ PRIOR FILING DATE: 2000-03-23
/ NUMBER OF SEQ ID NOS: 43008
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 270
/ LENGTH: 637
/ TYPE: PRT
/ ORGANISM: DROSOPHILA
/ US-11-097-143-270

Query Match      44.0%; Score 1174.5; DB 6; Length 637;
Best Local Similarity 49.3%; Pred. No. 7.4e-92;
Matches 245; Conservative 77; Mismatches 146; Indels 29; Gaps 9;
```

```
16 REDESEDESDLESPGCRKQKREOVNQGMPGLQSTFLAMDTEGVEVVMNLHG 75
84 RESGSDSESELESPGCRKWKREBEVDQRPGLDCVHLMDEGEVVMNVSQVA 143
76 DRKAPAAHEBKIQTFVEQVLVDHPNIVKLKHYMLDTSEA-CARVIFITEVSSGSLKOF 134
144 SIQELKSQEBKQKQVDFDNLQDHQNIKFKHYMTDTQABRRVVFITEVSSGSLKOF 203
135 LKTKKNKAMNARAWKRWCTQILSALSFILACSPPIHGNLTSDTIFIOHNGLIKIGSV 194
204 LKTKKNARLPLESWRMCTQILSALSYLHSCSPPIHGNLTSDTIFIOHNGLIKIGSV 263
195 WHRIFENALRPALDDELSPFRAERELRNHFF-PPRYGVADGT-AVDIFSGMCA 252
264 -----VPAVHYSVRGRERERERGAHYFOAPEYGAADLTALDLYAAGMCA 313
253 LEMAVLEIQTNQ--DTRVTEEAJARARHSISDPNMEFFILCLADPARRPSASHLFFHR 310
314 LEMAALEIQPSNESSTAINETIQRITFSLNDLQDLRKLKCNPPQDRPSANDLLFFHR 373
311 VLPEVHSLKLLAHCFTIOHYLMPENVVEKTKAMD-----LHAVLAELPRRRPPL 362
374 LTFEYHSLKLLTANHCIV---FSPANRTMSEFAFDGLMORYQPVVMAQLLAGGQR 429
363 OMRYSVSFMELDKLEEDVRNGIYPLMFAATRPGLRVLAAPPEEVOKAKTTPPEPP 421
430 QYRLAVDSGADKLEKFEVDKYGVPILYSGKKPNF-RSRAASPERADSVYSATPEV 488
422 DSETRKVIQMCNLERSEDKARWHLTLVLVEDRLRQLTYYDLPTDSADLASELVHYG 481
489 DRESRIIVMMCSVKIKEDSNDITWTITLMDKMKRQITCVQNMENDTAAIDLSELVRIG 548
482 FLHEDDRMKLAFLSEST 498
549 FVHLDDQDKIQVLLERT 565
```

```
RESULT 13
US-09-764-868-798
/ Sequence 798, Application US/09764868
/ Patent No. US2002016871A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: PT232
/ CURRENT APPLICATION NUMBER: US/09/764,868
/ CURRENT FILING DATE: 2001-01-17
/ Prior application data removed - refer to PALM or file wrapper
/ NUMBER OF SEQ ID NOS: 1510
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 798
/ LENGTH: 340
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SITE
/ LOCATION: (274)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (289)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (296)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (307)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (319)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ US-09-764-868-798

Query Match      35.4%; Score 946; DB 3; Length 340;
Best Local Similarity 60.4%; Pred. No. 1.4e-72;
Matches 201; Conservative 29; Mismatches 71; Indels 32; Gaps 7;
```

```
117 ARVIFITEVSSGSLKQPLKTKKNKANARAWKRWCTQILSALSFILACSPPIHGNL 176
5 AVIFITEVSSGSLKQPLKTKKNKANARAWKRWCTQILSALSYLHSCSPPIHGNL 64
177 TSDTIFIOHNGLIKIGSVWHRIFSNALRPPLALPDLSPIRAERELRNHFFPEYGE 236
65 TCDTIFIOHNGLIKIGSV-----APDTINNHVTCREQGNLHFFPEYGE 110
237 VAD-GTAVDIFSGMCALEMAVLEIQTNQDTR-VTEEAJARARHSISDPNMEFFILCCLA 294
111 VTNVTTAVDIYSGMCALEMAVLEIQNGESSYVPEALISSAIOLEEDPLQREFFIOKQ 170
295 RDPARRPSASHLFFRVLEPVHSLKLLAHCFTIOHYLMPENVVEKTKAMDHLAVLAEL 354
171 SEPARRPTRELLFHALFEPVSLKLLAHCIVGQHMPENALBEITOMDTSANVLAI 230
355 PR-PRRPPQWRYSEVSFEMLDKFLADVNGIYPLMFAATRPGLRVLAAPPEEVOKA 413
231 PAGGSEBPVQTLYSQSPALELDKFLADVNGIYPLTA-----GLPRXQOQOBEVTP 284
414 KTPPEPPDSETRKVIQMCNLERSEDKARWHL 446
285 VVP-EPVXDQFHTYSLRWL-----ARWXL 308

RESULT 14
US-09-764-875-1053
/ Sequence 1053, Application US/09764875
/ Patent No. US20040018969A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
/ FILE REFERENCE: P202
```

```

CURRENT APPLICATION NUMBER: US/09/764,875
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 1249
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1053
LENGTH: 340
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (274)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (289)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (326)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (307)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (319)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
IS-09-764-875-1053

```

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 12, 2006, 11:22:31 ; Search time 31 Seconds

(without alignments)
154.622 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAAPBPAPRRARERERED.....RMKLAAPLESTFLKRGTA 507

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 67062 seqs, 9454214 residues

Total number of hits satisfying chosen parameters: 67062

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: Published Applications AA New:
2: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB pep:
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB pep:
4: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB pep:
5: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB pep:
6: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB pep:
7: /cgn2_6/prodata/1/pubppa/US11_NEW_PUB pep:
8: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	199	7.5	588	6 US-10-821-234-1137	Sequence 1137, App
2	192	7.2	366	6 US-10-661-426-7	Sequence 7, App11
3	191	7.2	585	7 US-11-114-563-16	Sequence 16, App1
4	190	7.1	366	6 US-10-661-426-4	Sequence 4, App11
5	190	7.1	366	6 US-10-661-426-13	Sequence 13, App1
6	184.5	6.9	524	6 US-10-678-556A-131	Sequence 131, App
7	184	6.9	591	6 US-10-770-726-12	Sequence 71, App1
8	170.5	6.4	348	6 US-10-661-426-12	Sequence 12, App1
9	168.5	6.3	403	6 US-10-523-477-11	Sequence 11, App1
10	168.5	6.3	403	6 US-10-770-726-83	Sequence 83, App1
11	168.5	6.3	403	6 US-10-770-726-85	Sequence 85, App1
12	168.5	6.3	403	7 US-11-109-156-12	Sequence 12, App1
13	168.5	6.3	403	7 US-11-092-168-1	Sequence 12, App1
14	167.5	6.3	348	6 US-10-661-426-9	Sequence 9, App11
15	167.5	6.3	348	6 US-10-661-426-14	Sequence 14, App1
16	163	6.1	256	6 US-11-113-424-183	Sequence 183, App
17	157	5.9	1302	7 US-11-004-057-6	Sequence 6, App11
18	154	5.8	672	7 US-11-004-057-2	Sequence 2, App11
19	154	5.8	1493	7 US-11-004-057-4	Sequence 4, App11
20	153	5.7	277	7 US-11-151-601-4	Sequence 4, App11
21	151	5.7	1493	7 US-11-004-057-21	Sequence 21, App11
22	149	5.6	278	6 US-10-055-877-149	Sequence 149, App
23	149	5.6	278	7 US-11-103-065-4	Sequence 10, App1
24	149	5.6	278	7 US-11-151-601-10	Sequence 108, App
25	145	5.4	664	6 US-10-485-517-308	

26	142	5.3	712	6 US-10-770-726-66	Sequence 66, App1
27	140.5	5.3	970	6 US-10-523-477-10	Sequence 10, App1
28	140.5	5.3	970	6 US-10-770-726-86	Sequence 86, App1
29	140	5.2	547	6 US-10-770-726-87	Sequence 87, App1
30	139.5	5.2	532	6 US-10-860-501-7	Sequence 7, App11
31	138.5	5.2	297	6 US-10-770-726-48	Sequence 48, App1
32	138.5	5.2	297	7 US-11-109-156-11	Sequence 11, App1
33	138	5.2	445	6 US-10-770-726-70	Sequence 70, App1
34	137	5.1	656	6 US-10-821-234-1121	Sequence 1121, App
35	135.5	5.1	456	6 US-10-860-501-5	Sequence 5, App11
36	135.5	5.1	500	6 US-10-860-501-4	Sequence 4, App11
37	133.5	5.0	231	7 US-11-151-601-5	Sequence 5, App11
38	133.5	5.0	231	7 US-11-151-601-12	Sequence 12, App1
39	133.5	5.0	1311	6 US-10-509-422-5	Sequence 5, App11
40	131	4.9	298	6 US-10-770-726-51	Sequence 51, App11
41	130.5	4.9	418	7 US-11-099-691-1	Sequence 1, App11
42	129.5	4.9	418	7 US-11-109-156-2	Sequence 2, App11
43	129.5	4.9	1236	7 US-11-115-086-4	Sequence 4, App11
44	129.5	4.9	1306	6 US-10-995-561-905	Sequence 905, App
45	127	4.8	1061	7 US-11-059-814-18	Sequence 18, App1

ALIGNMENTS

RESULT 1	US-10-821-234-1137
Sequence 1137, Application US/10821234	
Publication No. US20050255114A1	
GENERAL INFORMATION:	
APPLICANT: Labat, Ivan	
APPLICANT: Seache-Crain, Birgit	
APPLICANT: Andarmant, Susan	
APPLICANT: Tang, Y. Tom	
TITLE OF INVENTION: Methode for Diagnosis and Treatment of Preeclampsia	
FILE REFERENCE: 821A	
CURRENT APPLICATION NUMBER: US/10/821,234	
CURRENT FILING DATE: 2004-04-07	
PRIOR APPLICATION NUMBER: US 60/462,047	
PRIOR FILING DATE: 2003-04-07	
NUMBER OF SEQ ID NOS: 1704	
SOFTWARE: pc_seq_genes Version 1.0	
SEQ ID NO 1137	
LENGTH: 588	
TYPE: PRT	
ORGANISM: Homo sapiens	
US-10-821-234-1137	
Query Match	7.5%; Score 199; DB 6; Length 588;
Best Local Similarity	20.6%; Pred. No. 7e-10;
Matches 125; Conservative 87; Mismatches 213; Indels 182; Gaps 27;	
QY	5 EPAPRR-----ARERERREDESEDESD-----ILRESPC 34
DB	9 DPAPRRPRPGCGSETTGGAAAAAAGGCGTGTGSESEFEGGPRAAAAMSDSSA 68
QY	35 GRWQKR-----EQVNOGMPGLOSTFLAMDTEGVEVWELHFGDKKAFPAHEKIQ 88
DB	69 LPMSIRRDYELQEVIGSATAVQAACAPKKE---KVAIKRINL-----EKCO 115
QY	89 TYPEDQV-----LVDHENIVKHKYMDTSEACARVFTREYVSGSLQKPLK--TK 139
DB	116 TSNDELKKEIQAMSGCHPNIVSYYSFVAKDE---LWLWKLSSGSGVLDIKHIVAK 171
QY	140 KHKKA--NNARAKWKCQILSLSPFIHGNLTSPDTFIHNGLIKIGSMHR 197
DB	172 GHHKSVLDESTIATITREVLGLEYLHNGQ--IHRDVKAGNIILGEGSVQIADFGVS 229
QY	198 IFSNALRPPTALPDDIRSPTRAREELRNLIH-----FPPPEYGEVADG--TAVDIFSPGM 250
DB	230 AF-----LATGCDI-----TRNKVTRKTFVGTGPMAPMEVMEQVRYGYDFRADISFGSI 276
QY	251 CALDMA-----VLEIQTNQDTRVTEBAIARARHSLSDPM-----RETI 289

```

Db      277 TAIELATGAAPYHKYDPMKVLMTLQNDPESLETGV-----QDKEMLKVKYKSPKRMII 329
Qy      290 LCLLADPARPBAHSHLLFHRVLFVYHSLKLLAAHCFIQHYLMPEVNVBEKTKAMDILHA 349
Db      330 SLCLODPEKRPFAELHHRKFFQAKNKE-----FLOEKTLOQRAPTTSERAK----- 377
Qy      350 VLAELEPRPRPPLQ-----WRYSEVSPMEIDKFLVDVNRNGIYPLAN----- 390
Db      378 -----KVRVPRSSGRLLKTEBDGWMWDDER---DESEBCKAIIQSRLSPRVSEIS 428
Qy      391 ----FAATPPLGLPRVLAPRPEEV-----QKATPEPPEPSETKRVIQ 430
Db      429 NSELFPETDPVG---TLQVPEQISAHLPQAPQIATOPTQVSLPPTAPR--AKTAQALS 483
Qy      431 MGCNLERSEDKAMNHLLVLVLEDRHRLQTYDLLP-TSADQLAEVLYHGFELHDDDM 489
Db      484 SSGSGQ--ETKLPISLVLRNRSKKELANDIRFEFTGRTDPAEGVSGELISAGLVDRDLV 541
Qy      490 KLAAPLE 496
Db      542 IVANAIQ 548

RESULT 2
US-10-661-426-7
; Sequence 7, Application US/10661426
; Publication No. US20050262584A1
; GENERAL INFORMATION:
; APPLICANT: Sneed, Jen
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Asai, Tsuneaki
; APPLICANT: Tena, Guillaume
; TITLE OF INVENTION: Master Activators of Pathogen Responsive
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 00786/397003
; CURRENT APPLICATION NUMBER: US/10/661,426
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: PCT/US 02/07650
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/275,199
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FaestSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-661-426-7

Query Match      7.2%; Score 192; DB 6; Length 366;
Best Local Similarity 22.6%; Pred. No. 1.5e-09;
Matches 82; Conservative 53; Mismatches 122; Indels 106; Gaps 16;

Qy      45 NOGNMGLQSTFLAMDTBEGVE-----VVMNELHFGDKKAFPA--- 82
Db      51 SSGSAASSGSGSASTTNSIEAKVNSDLVRGNRISGAGTYVKKYIHRPSSRLVYLKVI 110
Qy      83 ---HEKTIQ---TVBEQVLVDHENIVKLHKYMTLDTSEACARVIFITTVYSSGSLKQPL 135
Db      111 YGNHEETVRRQICREIEILRDVNHVNVKCHEMFQNGE---IQVLLEFMDKGSLE--- 163
Qy      136 KTKTKKHKKAMNANAKMR-----WCTQILSLSLFHACSPPIIHGNTSTPTIFIQHNGULK 190
Db      164 -----GAHVMEQQLADLSRQILSGLAYLH--SHYIVRDIKPSNLLINSKAVK 211
Qy      191 IGSVW--HRIFSNALRPPALPDDLRSPIDAREBELNLHFFPPE-----YGEVADGTA 242
Db      212 IADFGVSRILAQMDPC-----NESVGTIAYMSPERINTDLNGCKY-DGTA 256
Qy      243 VDIFFSGMCALEMAV---LEIQTNGDTRVTEBAIARAR---HSLSDPNMEFFILCCLA 294
Db      257 GDWISIGVILBEFLGRFPFPVSRQGDWASIMCAICMSQPEAPATVASEPFRHFISCIQ 316

```

```

Qy      295 RDPARPSAHSLLFHRVLFVYHSLKLLAAHCFIQHYLMPEVNVBEKTKAMDILHA 354
Db      317 REFGKRRAMQILL-----QHPFIRASPSQNR--PQNLHQULPP- 354
Qy      355 PRP 357
Db      355 PRP 357

RESULT 3
US-11-134-563-16
; Sequence 16, Application US/11134563
; Publication No. US20050287569A1
; GENERAL INFORMATION:
; APPLICANT: Leong, John M.
; APPLICANT: Campellone, Kenneth G.
; TITLE OF INVENTION: ESFPU NUCLEIC ACIDS AND PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 07917-280001
; CURRENT APPLICATION NUMBER: US/11/134,563
; CURRENT FILING DATE: 2005-05-20
; PRIOR APPLICATION NUMBER: US 60/573,600
; PRIOR FILING DATE: 2004-05-20
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FaestSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-134-563-16

Query Match      7.2%; Score 191; DB 7; Length 545;
Best Local Similarity 24.6%; Pred. No. 3.2e-09;
Matches 84; Conservative 60; Mismatches 115; Indels 82; Gaps 16;

Qy      4 PEPARPRAREREREREDESEDSILIE-----SPCGWQKREQVNOGNMGLQST 55
Db      221 PPDALTTRMTEKQKKPKKPSDSE--ILEKLRSIVSGDRKKKY--TRFKIKGQ---ASGTV 284
Qy      56 FLAMDTBEGVEVVMNELHFGDKKAFPAHBEKIQVFEQVLVY--DHNIVYKLHKYMLDT 112
Db      285 YTAMDVATGOEVAIQQNML-----QQQPKKELIINELVWRKNKNPNIYA---YIDS 333
Qy      113 SEACARVIFITEYVSSGSLKQFLKTKKHKKAMNANAKMRCTQILSLSLFHACSPPI 172
Db      334 YLVGDELWVMEYLAGSSLTDVYET-----CMDEGQIAAYCRBCLQALBEFLH--SNQVI 386
Qy      173 HGNLTSDPTFIQHNGLIKIGSVWHRIFSNALRPPALPDDLRSPIDAREBELNLHFFP-- 230
Db      387 HRDISDNIILGMDQSVLTDPRG---FCAQI-----TPQSKRSTVNGTPYMYA 432
Qy      221 PPEYGEVADGTAVDIFSGMCALEM-----AVLEIQTNGDTRVTEBAIARA 276
Db      433 PEVVTAKAYGPKVDIWSIGIWAIEWIEGEPYLNENPRLAYLLIATNG---TPE----- 483
Qy      277 RHSUSDPN-----WBEFILCLLADPARPBAHSHLLFHRV 312
Db      484 ---LQNPFLKSAIFRDFILNRCLMDVERKRGSAKELLQHOFL 521

```

```

RESULT 4
US-10-661-426-4
; Sequence 4, Application US/10661426
; Publication No. US20050262584A1
; GENERAL INFORMATION:
; APPLICANT: Sneed, Jen
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Tena, Guillaume
; APPLICANT: Asai, Tsuneaki
; TITLE OF INVENTION: Master Activators of Pathogen Responsive
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 00786/397003

```



```

CURRENT APPLICATION NUMBER: US/10/661,426
CURRENT FILING DATE: 2003-09-12
PRIOR APPLICATION NUMBER: PCT/US 02/07650
PRIOR FILING DATE: 2002-03-13
PRIOR APPLICATION NUMBER: US 60/275,199
PRIOR FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 366
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-661-426-4

```

Query Match	7.1%;	Score 190;	DB 6;	Length 366;
Best Local Similarity	23.3%;	Pred. No. 2.2e-09;		
Matches	81;	Conservative	54;	Mismatches 122;
				Indels 106;
				Gaps 16;

```

QY      45 NOONMGLSTFIAMDTEEGVE-----VWNHELFGGRKAFPA--- 82
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      51 SSSSAAPSSGGSASTTNNSSIEAKNYSDLVNRNIGSAGGCVTVYVTHRPSRLYALKVY 110
      : : : : : : : : : : : : : : : : : : : : : : : :
QY      83 ---HEEKIQ---TFEOLVLDHPNIVYKHLKMYLDTSEACARVIFITEYVSSGLKQFL 135
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      111 YGNHEETVRQICREIETILRDVNHPRVYVCHEMPONGE---IQLLEFMDKSL- 163
      : : : : : : : : : : : : : : : : : : : : : : : :
QY      136 KTKKQHKHKKMMAPAKWR-----WCTOILSALSFLLACSEPIITHGNLVSDTIFIOHGLIK 190
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      164 -----GAHVWEEOQLADLSROLISGLAYLH--SRHLYHRDIKPSNLLINSARVYK 211
      : : : : : : : : : : : : : : : : : : : : : : : :
QY      191 IGSWV-HRIFSNLRPEPTALPDOLNSPIABREELRNHFFRPE-----YGRVADGTA 242
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      212 IADFGYSRIIAQTMDCNS-----SVGTATYSPERINTDNLQGRY-DGYA 256
      : : : : : : : : : : : : : : : : : : : : : : : :
QY      243 VDIISFGMCALENAV---LEIQOTGDFRVTETEARAR---HSLSDPNMRREFTICLA 294
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      257 GDIIWSIGVSTILBEYTLGRPPFPVPSROGDWASLCAICMSQPEAPATPSPEBRHHIISCLQ 316
      : : : : : : : : : : : : : : : : : : : : : : : :
QY      295 RDPARRPSAHSLLFHRVLEFVHSLKLLAAHCTIQOYIMPENVEEKTAKMDLHAYLAEL 354
      : : : : : : : : : : : : : : : : : : : : : : : :
Db      317 REFGKSRASQML-----QHPFILRASPSQONS-PONLHQLLP- 354
      : : : : : : : : : : : : : : : : : : : : : : : :
QY      355 PRP 357
      : : : : :
Db      355 PRP 357

```

```

US-10-661-426-13
/ RESULT 5
/ US-10-661-426-13
/ Sequence 13, Application US/10661426
/ Publication No. US20050262584A1
/ GENERAL INFORMATION:
/ APPLICANT: Sheen, Jen
/ APPLICANT: Ausubel, Frederick M.
/ APPLICANT: Aaal, Teuneaki
/ APPLICANT: Tena, Guillaume
/ TITLE OF INVENTION: Master Activators of Pathogen Responsive
/ TITLE OF INVENTION: Genes
/ FILE REFERENCE: 00786/397003
/ CURRENT APPLICATION NUMBER: US/10/661,426
/ CURRENT FILING DATE: 2003-09-12
/ PRIOR APPLICATION NUMBER: PCT/US 02/07650
/ PRIOR FILING DATE: 2002-03-13
/ PRIOR APPLICATION NUMBER: US 60/275,199
/ PRIOR FILING DATE: 2001-03-12
/ NUMBER OF SEQ ID NOS: 16
/ SOFTWARE: FastSeq for windows version 4.0
/ SEQ ID NO 13
/ LENGTH: 366
/ TYPE: PRF
/ ORGANISM: Arabidopsis thaliana
/ US-10-661-426-13

```

Query Match	7.1%;	Score 190;	DB 6;	Length 366;
Best Local Similarity	22.3%;	Pred. No. 2.2e-09;		
Matches	81;	Conservative	54;	Mismatches 122;
				Indels 106;
				Gaps 16;

```

QY      45 NOQNMGLSTFLAMDTEGVE-----VWNELLFGRKAFPA--- 82
D      51 SSSSAASSSGSSASSNTTNSITAKRYSDLYKGNRTGSGAGGTVKVTHREPSRIYALKVI 110
QY      83 ---HEEKIO---TYPEQLVLDVHDNIYKHLKMYLDTSEACARVIFTTEYVSSGSLQFL 135
D      111 YGNHEETVARQICREIEILRDVHNHNHYVKCHMFDPNGE-----IQVLLEFMDKGSLE--- 165
QY      136 KTKKKNHKKMNRAMWR-----WCTQIISALSFLHACSPPIIHGNLSDTFTFIOHNGLIK 196
D      164 -----GHWKKEOQLADLSRQIISGLAYLH--SRHIVHDIKPSNLLINSANVKV 211
QY      191 IGSWV-HRIFSNALRPPTALPDLLSRPIRAREBELRNLHFEPPE-----YGEVADGTA 244
D      212 IADFGVSRLLAQTMPCNS-----SVGTATYMSPERINTDLNGGX-DGYA 256
QY      243 VDIFFSGMCALEMAV---LBIOTNGDRTVTEBAIARAR-----HSLSDPNMRREFITCCLA 294
D      257 GDIWLSGVSIIEFYIGRPFPPVRSQGDWASLMCAI CMSQPEAPATASPEPRHHIISCCLO 316
QY      295 RDPARPRPSASHLLFHRVLFVYHSLKTLAHCPIQHQYIMPENNVVEKTKAMDHLAVIAEL 354
D      317 REBGKKRSAMOLL-----OHPFILRASPEQNRNS-PONILHQLLP- 354
QY      355 PRP 357
D      355 PRP 357

```

```

US-RESULT 6 RESULT-6-556A-131
/ Sequence 131, Application US/10878556A
/ Publication No. US20050266399A1
/ GENERAL INFORMATION:
/ APPLICANT: Hoffmann La-Roche Inc.
/ TITLE OF INVENTION: HCV regulated protein exp
/ FILE REFERENCE: 21762
/ CURRENT APPLICATION NUMBER: US/10/0,878,556A
/ CURRENT FILING DATE: 2004-06-28
/ NUMBER OF SEQ ID NOS: 199
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 131
/ LENGTH: 524
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: gw_hum/pak2_humam
/ DATABASE ENTRY DATE: 1997-11-01
/ US-10-878-556A-131

```

Query Match	6.9%;	Score 184.5;	DB 6;	Length 524;
Best Local Similarity	23.3%;	Pred. No. 1.1e-08;		
Matches	80;	Conservative	60;	Mismatches 126;
			Indels	77;
			Gaps	15;

QY 4PPAP-----RRRRERERREDESBESJILTE-----SPGRMOKREBOVQNG 48
 Db 201 PPAFAVGVSHDGAANKSLDKOKKPKMTDEIMTKATTIYSIGDPKKKY-TRIYKIGQG- 258
 QY 49 MEGLOSTFLAMDTBEGVEVWNNEILHFGDKAPAAHEKIQTVEQVLV---VHPNIVKL 105
 Db 259 --ASGVFTATDVNALGQEVALKQINL-----QKPKKELINELWMLKNPNIVN- 308
 QY 106 HKTWLDTSBACARVITFTEYVSSGSLKQFLKTKKHKANAPARWKCTQIILSALSPH 165
 Db 309 ---FLDSYLVGDELFTVMVMEYLAGSLDVTVET-----CMDDAQIAAVCRECQIATLEFHL 360
 QY 166 ACSPIIITGNLTSDFITFIQHNGLKIGSVMHRIFSNMLRPTPALPDOLRSPFAAREBELR 225
 Db 361 A--NOVIRHDKISDVLNLAGMEGSVYKLDFG---SCAQI-----TPROSKSTWV 404

Qy 226 NLHFF-PRVEGEVLDGTVDI-PSPEMCALEM-----AYLEIOMNGDTRVY 269
 :::::| | | | |
 Db 405 GTFYMAPEVTVTKAYGPKVDIMSLGIMILEWGESEPTLYINMPALNYLIATNGTPELO 464
 :::::| | | | |
 Qy 270 BEAIPARHSLSDDPNMKREPTLCCLARDPAPRPSAHSLLFRRVY 312
 :::::| | | | |
 Db 465 NP-----EKUS-PIFRDPLNRCLENDVKEGSAKELLCHPFL 500
 :::::| | | | |

RESULT 7
US-10-7770-726-71

```

RESULT 8
US-10-661-426-12
; Sequence 12, Application US/10661426
; Publication No. US20050262584A1
; GENERAL INFORMATION:
; APPLICANT: Sheen, Jen
; APPLICANT: Ausubel, Frederick M.
; APPLICANT: Asai, Tsuneaki
; APPLICANT: Tena, Guillaume
; TITLE OF INVENTION: Master Activators of Pathogen Responsive
; TITLE OF INVENTION: Genes
; FILE REFERENCE: 00786/397003
; CURRENT APPLICATION NUMBER: US/10/661,426

```

```

; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: PCT/US 02/07650
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/275,199
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 12
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-661-426-12

```

Query Match	6.4%	Score 170.5;	DB 6;	Length 348;
Best Local Similarity	24.5%;	Pred. No. 1.1e-07;		
Matches 67;	Conservative 38;	Mismatches 103;	Indels 65;	Gaps 12;

```

RESULT 9
US-10-523-477-11
/ Sequence 11, Application US/10523477
/ Publication No. US20050266406A1
/ GENERAL INFORMATION:
/ APPLICANT: EXELIXIS, INC.
/ TITLE OF INVENTION: MAXS AS MODIFIERS OF THE AXIN PATHWAY AND METHODS OF USE
/ FILE REFERENCE: EX03-051C-US
/ CURRENT APPLICATION NUMBER: US/10/523,477
/ CURRENT FILING DATE: 2005-02-06
/ PRIOR APPLICATION NUMBER: US 60/401,534
/ PRIOR FILING DATE: 2002-08-07
/ PRIOR APPLICATION NUMBER: US 60/411,153
/ PRIOR FILING DATE: 2002-09-16
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 11
/ LENGTH: 403
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-523-477-11

```

Query Match	6.3%	Score 168.5	DB 6	Length 403
Best Local Similarity	24.0%	Pred. No. 2e-07		
Matches	76	Conservative 61	Mismatches 131	Indels 49
			Gaps 14	
QY	2	AAPBPAPRARERE---	REREDESEDSDLIEESPCCRMQKRGVQGNNGNAPGLQSTPLA	58
		:::	:::	:::
DB	104	SAPENNP---EEELASQKQKESSKKQOMALEDEPIGR----	PLGKKGK--FGNVYLLA	150
QY	59	MDTEEGEVVWVNEIHFQGDRAKFAAHEEKIQVFEEQLVIVHPNIVKLTKYVLDTSBACAR		118
	:::	:::	:::	:::
DB	151	REKSKFTLAKVL-FKAQLEKAGVEHQRRREVETIQSHLRPNILRLRYGYHDAT----	R	205
QY	119	VIFTEVVSQSLKQPLKTKKGNKAMNARAWKWCITQIISALSPFLHACSPDIHGNITS		178


```

Db      8  VYKVIHPTSPALPAKVITIGNEDTVRQICREIELLSVDHPNVKCH-----DMFPHNG 133
Qy      118 RVLPFTFYSSGSLKQFLPKTKKKNKAMAMKMR-----WCTQLSLSLFLACSPPII 172
Db      140 EIVULLLEFMQGLE-----GAIHWOBQELADLSRQTLISGLATIH--RRHIV 184
Qy      173 HGNLTSDPTPIQHNGLIKIGSVW-HRIFSNALRPPALPDDLSPTRAREELNHLHFP 231
Db      185 HRDIKPSNLLINSAKNVKIADPROVSKILLAQTMDCNS-----SVGTIYMS 230
Qy      232 PE-----YGEVADSTAVDIPSGFCALEMAV-----LEIQNGDTRVTEEARAR-- 277
Db      231 PERINTDNLNGRY-DEYAGDVMISLGILEFYLGFRPFAVSQGDGMASLMCAIMCSQPE 289
Qy      278 -HSLSDPMNEFLTCLCLADPPARPRPAHSLHLTH 309
Db      290 APATASOEFRHFVSCCLQSDPPKRMASQAQQLQH 322

```

RESULT 15

```

US-10-661-426-14
; Sequence 14, Application US/106611426
; Publication No. US20050262584A1
GENERAL INFORMATION:
APPLICANT: Sheen, Jen
APPLICANT: Ausubel, Frederick M.
APPLICANT: Asai, Tsuneaki
APPLICANT: Tena, Guillaume
TITLE OF INVENTION: Master Activators of Pathogen Responsive
TITLE OF INVENTION: Genes
FILE REFERENCE: 00786/397003
CURRENT APPLICATION NUMBER: US/10/661,426
CURRENT FILING DATE: 2003-09-12
PRIORITY APPLICATION NUMBER: PCT/US 02/07650
PRIORITY FILING DATE: 2002-03-13
PRIORITY APPLICATION NUMBER: US 62/275,199
PRIORITY FILING DATE: 2001-03-12
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14
LENGTH: 348
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-661-426-14

```

Query Match	6.3%	Score 167.5;	DB 6;	Length 348;
Best Local Similarity	24.2%;	Pred. No. 2e-07;		
Matches 66;	Conservative 39;	Mismatches 103;	Indels 65;	Gaps 12;

```

QY 68 VNNELIFGRRKAPAA-----HEEKIQ-----TFEGQLYLDHPNIVGLHKHYMLDTSBAC 117
Db 84 VYKVVHTTPSRPALKNVLYGNHEDTVKROICREIFILSVDPHNVVKCH---DMFHNG 139
QY 118 RVLFITEVYSSGLKQFLKTKRKHKANAMARMK-----WCTQILSALFPLHACSPRII 172
Db 140 EIQVLLEFMDQSLF-----GAIHWGQELADLRQILSLGLAYLH--RRHIV 184
QY 173 HGNLTSDTIFIQHNGLIKIGSVW--HRIFSNLRPTPLRPDLRSPIRAEKEELRNHLFFP 231
Db 185 HRRIKPSNLLINSANKVKIADFGVSRILIAQTMDPCNS-----SVGTAYMS 230
QY 232 PE-----YGEVADSTAVDIFSPFCALENAV-----LEIQNGDTRVTEEARAR--- 277
Db 231 PEKINTDLMHGRV--DSYAGDVMSLGVSLIEPLFGFPFAVSRQGDMAIWCALCMQSPPE 289
QY 278 -HSLSPNNRBEPLCTIANDPARPSANSLFLH 309
Db 290 AAPTASQAEFRNHFVSCLOSDEPKTASAOQLQH 322

```

Search completed: January 12, 2006, 11:33:39
Job time : 32 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 14, 2006, 00:46:17 ; Search time 385 Seconds
(without alignments)
1065.817 Million cell updates/sec

Title: US-10-618-941-99
Perfect score: 2670
Sequence: 1 MAAPAPAPRARERERERED.....RMKLAFLSTLKYKRGTOA 507

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Zgapop 6.0, Zgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6038814 seqs, 404674181 residues
Total number of hits satisfying chosen parameters: 12077628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame.p2n.model -DEV=xlp
-q=/cgn2_1/USPTO.spool_p/US10618941/runac_12012006_070207_14416/app_query.fasta_1.647
-DB=Published Applications NA New -OPMT=fastap -SUFPIX=rnpbn -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=apct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=local -OUTPMT=pro -NOR=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10618941@CGN_1_1_184_@runac_12012006_070207_14416
-NCPUB=1 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA New.*

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq2:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq3:*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	413	15.5	540	7	US-11-128-061-2970
2	413	15.5	540	7	US-11-128-061-6612
3	383.5	13.5	669	6	US-10-623-155-23
4	359.5	13.5	6812	6	US-10-953-054-44
5	302.5	11.3	7555	7	US-11-136-527-688
6	199	7.5	2838	6	US-10-770-726-27
7	199	7.5	4529	6	US-10-821-234-285
8	192	7.2	1101	6	US-10-661-426-6

9	191	7.2	1400	7	US-11-136-527-6476	Sequence 6476, Ap
10	191	7.2	1638	7	US-11-134-563-15	Sequence 15, Appl
11	191	7.2	2539	7	US-11-136-527-2380	Sequence 2380, Ap
12	190	7.1	1101	6	US-10-661-426-3	Sequence 3, Appl1
13	185.5	6.9	1964	7	US-11-136-527-3130	Sequence 3130, Ap
14	180	6.7	2241	7	US-11-136-527-2475	Sequence 2475, Ap
15	173	6.5	2253	6	US-10-770-726-39	Sequence 39, Appl
16	173	6.5	2253	7	US-11-000-688-124	Sequence 124, App
17	173	6.5	2347	6	US-10-770-726-41	Sequence 41, Appl
18	172.5	6.5	2033	6	US-10-523-477-4	Sequence 4, Appl1
19	170.5	6.4	1047	6	US-10-661-426-11	Sequence 11, Appl1
20	167.5	6.3	1047	6	US-10-661-426-8	Sequence 8, Appl1
21	166	6.2	2605	7	US-11-000-688-1480	Sequence 1480, Ap
22	157	5.9	3911	9	US-11-004-057-5	Sequence 5, Appl1
23	154	5.8	3260	9	US-11-004-057-1	Sequence 1, Appl1
24	154	5.8	5253	9	US-11-004-057-3	Sequence 3, Appl1
25	150	5.6	3845	7	US-11-136-527-2764	Sequence 2764, Ap
26	148	5.5	4453	7	US-11-000-688-1248	Sequence 1248, Ap
27	147.5	5.5	3293	6	US-10-770-726-43	Sequence 43, Appl
28	145	5.4	2248	6	US-10-485-517-84	Sequence 84, Appl
29	145	5.4	3258	7	US-11-136-527-3018	Sequence 3018, Ap
30	144.5	5.4	1891	6	US-10-860-501-2	Sequence 2, Appl1
31	144.5	5.4	2056	6	US-10-860-501-1	Sequence 1, Appl1
32	143.5	5.4	4554	7	US-11-136-527-2255	Sequence 255, App
33	142.5	5.3	3837	7	US-11-136-527-90	Sequence 90, Appl
34	142	5.3	3583	6	US-10-770-726-22	Sequence 22, Appl
35	140.5	5.3	1327	6	US-10-523-477-3	Sequence 3, Appl1
36	140.5	5.3	3331	6	US-10-770-726-42	Sequence 42, Appl
37	140	5.2	3676	7	US-11-136-527-3587	Sequence 3587, Ap
38	139.5	5.2	2932	6	US-10-860-501-8	Sequence 8, Appl1
39	138.5	5.2	1235	6	US-10-770-726-4	Sequence 4, Appl1
40	138	5.2	2119	6	US-10-770-726-26	Sequence 26, Appl
41	138	5.2	2328	6	US-10-770-726-7	Sequence 7, Appl1
42	138	5.2	2328	7	US-11-000-688-154	Sequence 154, App
43	137.5	5.2	5464	7	US-11-136-527-3219	Sequence 3219, Ap
44	137.5	5.1	2572	7	US-11-136-527-2589	Sequence 2589, Ap
45	137.5	5.1	3527	6	US-10-770-726-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-11-128-061-2970
; Sequence 2970, Application US/11128061
; Publication No. US20060003958A1
; GENERAL INFORMATION:
; APPLICANT: Melville, Mark W.
; APPLICANT: Charlebois, Timothy S.
; APPLICANT: Mounts, William M.
; APPLICANT: Ham, Louane E.
; APPLICANT: Sinacore, Martin S.
; APPLICANT: Leonard, Mark W.
; APPLICANT: Brown, Eugene L.
; APPLICANT: Miller, Christopher P.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
; FILE REFERENCE: 01997.027701
; CURRENT APPLICATION NUMBER: US/11/128,061
; PRIOR FILING DATE: 2005-05-11
; PRIOR APPLICATION NUMBER: US 60/570,425
; NUMBER OF SEQ ID NOS: 7285
; SOFTWARE: PatentIn version 3.3
; LENGTH: 540
; TYPE: DNA
; ORGANISM: *Cricetulus griseus*
US-11-128-061-2970
Alignment Scores:
Pred. No.: 8, 226-37 Length: 540
Score: 413.00 Matches: 74
Percent Similarity: 81.82% Conservative: 16

```

Best Local Similarity: 67.27%      Mismatches: 20
Query Match: 15.47%              Indels: 0
DB: 7                             Gaps: 0

US-10-618-941-99 (1-507) x US-11-128-061-2970 (1-540)

QY      17  GUAAGGUAAPGUGUSERGUAAPGUSERAPPIELEUGUGUSERPROCYGGLYARG 36
      |||:::|||||
DB       211  GAGGAGGAAGAAAGTGTAGTGTGATCTTGAMAGAGTCCCTGTGACGC 270

QY      37  TTPGILYARGARGUGUGUVALAANGINGLYASMEPROGILYENGUSERTHPHE 56
      |||:::|||||
DB       271  TGGCAGAAAGGAGAGAGAGGTGAATCAGCCGTAATGACAGATTACAGTGCATAC 330

QY      57  LEUAIAMEAASPHTHRUGUGUGLYVALGIUVALVATTPAANGULUEHISPHGGLYASP 76
      |||:::|||||
DB       331  CTGGCCATGATATCAGAGGAGAGGTGTGGAGGTGTGTGGATAGAGACGTTCTCGAA 390

QY      77  ARGLYALPHEALAHISGUGUGULYBILGINTHRVALPHEGUGULUEVALLEU 96
      |||:::|||||
DB       391  CCCAAGAACTACCAACTGCAGGAGGAAAGGTCCGTCAGCTGTTGTAATTCATTCACAA 450

QY      97  VALASPHISPROASNTLEVALYLSLEAHISLYETRTPLEUNSPHTUSERGILUALCYE 116
      |||:::|||||
DB       451  TTGGAACTACTTACATCTGTTAAGTTTCCACAGATATGGGCTGATGTTAAAGAAATAA 510

QY      117  ALARGVALIIPHEIETHRUGIYRYVAL 126
      |||:::|||||
DB       511  GCCCGGTGATTTTCATTCACAGAAATCATG 540

RESULT 2
US-11-128-061-6612
/ Sequence 6612, Application US/11128061
/ Publication No. US20060003958A1
/ GENERAL INFORMATION:
/ APPLICANT: Melville, Mark W.
/ APPLICANT: Charlebois, Timothy S.
/ APPLICANT: Mounts, William M.
/ APPLICANT: Hann, Louane E.
/ APPLICANT: Sinacore, Martin S.
/ APPLICANT: Leonard, Mark W.
/ APPLICANT: Brown, Eugene L.
/ APPLICANT: Miller, Christopher P.
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS
/ FILE REFERENCE: 01997.027701
/ CURRENT APPLICATION NUMBER: US/11/128, 061
/ PRIOR FILING DATE: 2005-05-11
/ PRIOR APPLICATION NUMBER: US 60/570,425
/ NUMBER OF SEQ ID NOS: 7285
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 6612
/ LENGTH: 540
/ TYPE: DNA
/ ORGANISM: Cricetus griseus
US-11-128-061-6612

Alignment Scores:
Pred. No.: 8.22e-37      Length: 540
Score: 413.00           Matches: 74
Percent Similarity: 61.82%      Conservative: 16
Best Local Similarity: 67.27%      Mismatches: 20
Query Match: 15.47%          Indels: 0
DB: 7                   Gaps: 0

US-10-618-941-99 (1-507) x US-11-128-061-6612 (1-540)

QY      17  GIUARGLIAAPGUGUSERGUAAPGUGUSERAPPIELEUGUGUSERPROCYGGLYARG 36
      |||:::|||||
DB       211  GAGGAGGAAGAAAGTGTAGTGTGATCTTGAAATCTTGAAAGAGTCCCTGTGACGC 270

QY      37  TTPGILYARGARGUGUGUVALAANGINGLYASMEPROGILYENGUSERTHPHE 56
      |||:::|||||
DB       271  TGGCAGAAAGGAGAGAGGTGAATCAGCCGTAATGACAGATTACAGTGCATAC 330

QY      57  LEUAIAMEAASPHTHRUGUGUGLYVALGIUVALVATTPAANGULUEHISPHGGLYASP 76
      |||:::|||||
DB       331  CTGGCCATGATATCAGAGGAGAGGTGTGGAGGTGTGTGGATAGAGACGTTCTCGAA 390

QY      77  ARGLYALPHEALAHISGUGUGULYBILGINTHRVALPHEGUGULUEVALLEU 96
      |||:::|||||
DB       391  CCCAAGAACTACCAACTGCAGGAGGAAAGGTCCGTCAGCTGTTGTAATTCATTCACAA 450

QY      97  VALASPHISPROASNTLEVALYLSLEAHISLYETRTPLEUNSPHTUSERGILUALCYE 116
      |||:::|||||
DB       451  TTGGAACTACTTACATCTGTTAAGTTTCCACAGATATGGGCTGATGTTAAAGAAATAA 510

QY      117  ALARGVALIIPHEIETHRUGIYRYVAL 126
      |||:::|||||
DB       511  GCCCGGTGATTTTCATTCACAGAAATCATG 540

```

```

Db      271  TGGCAGAAAGAGAGAGAAAGGATGATCGCCCTAAATGTACAGAGTATTTGACAGTGTGATAC 350
Oy      57  LeuAlaMetAspThrGluGluGluValAlaTrpAsnGluLeuHisPheGlyAsp 76
Db      331  CTGGCGCATGTGATACAGAGAGAGAGGTGTGGAGGTGTGTGTGTGTGAAATACAGGTACTCTCGAA 390
Oy      77  ArgTyrAlaPheAlaAlaHisGluGluTyrIleGlnThrValPheGluGluLeuValLeu 96
Db      391  CGCAAGAACTACAAACTGCAGCAGGAGAAAGAGTCCCGTCGACAGTGTGTATATATCATTCATCAA 450
Oy      97  ValAspHisPheProAsnIleValLeuHisIleTyrTrpLeuAspThrSerGluValAsp 116
Db      451  TTGGAGACATCTTAAACATCGTTAAGTTTCAACAATATTTGGCTGATGTTTAAAGAAATTTAA 510
Oy      117  AlaArgValIlePheIleThrGluTyrVal 126
Db      511  GCCCGGTGATTTTCATCATCAGAAATACAG 540

RESULT 3
US-10-623-155-23
; Sequence 23, Application US/10623155
; Publication No. US20050261166A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongrong
; APPLICANT: Beckham, David W.
; APPLICANT: Retter, Marc W.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.45C20
; CURRENT APPLICATION NUMBER: US/10/623.155
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23
; LENGTH: 669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 642..661
; OTHER INFORMATION: n = A,T,C or G

US-10-623-155-23

Alignment Scores:
Pred. No.:      2,686-33      Length:      669
Score:          383.50      Matches:      83
Percent Similarity: 66.24%      Conservative: 21
Best Local Similarity: 52.87%      Mismatches:  21
Query Match:    14.36%      Indels:      13
DB:             6          Gaps:          3

US-10-618-941-99 (1-507) x US-10-623-155-23 (1-669)

Oy      349  AlaValLeuAlaGluLeuProArg--ProArgArgProProLeuGlnTrpArgTyrSer 367
Db      7   GCCGTACTGGCTGGAATATCCCTGCAGAGACCAAGAAAGAAACCAAGTTCAAGATTGTACTCT 66
Oy      368  GluValSerPheMetGluLeuAspLysPheLeuGluAspValArgAsnGlyIleTyrPro 387
Db      67  CAGTACACCAAGCTCTCGAAATTAAATTCTTTGAAAGATGTCAAGAAATGGAGACTATTCCT 126
Oy      388  LeuMetAsnPheAlaAlaThrArgProLeuGlyLeuProArgValLeuAlaProProPro 407
Db      127  CTGACAGCCCTT-----GGAGCTGCTTCGGGCCCAAGACCAAGCAG 168
Oy      408  GluGluVal-----GlnLysAlaLysThrProThrProGluProPhe 421
Db      169  GAGGAGGTGTGACATCACTGCTGTGCCCCCTCTGTCAAAAGCTCCGACACTGAAACCAAGCT 228
Oy      422  ApsSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLys 441

```


Dd	222	GAGGTGGACACTCGCAAGGTCGTGATGCATGCAACATTGAATCGCTGGAGAGGGA	288
Oy	442	A1aaagtTgHieLeuthrleuleuleuValleuglunapragleunihargjInLeuthr	461
Dd	289	GTCAAACAACCCTTGACACTTCTGTGTAAGTTGAGAGCAAACTGMACCGGCACCTGAGC	348
Oy	462	TyrApRleuLeuProthrXerSerAlagInApRleuAlaserGluleuValhistrYglY	481
Dd	349	TCGTACTGTATGCGAATAATGAGAATATCCCCGAGATGGCGCTGAGCTGTGACGTGGGC	408
Oy	482	PheLeuHieGlunApRApRmeClYleuAlalalePheleugIuserThr	498
Dd	409	TTCATTTAGTAGGCTGACCAAGCCGGTTGACTTCTCTGTTAGAGAGACT	459
 RESULT 4 US-10-955-054A-44 ; Sequence 44, Application US/10955054A ; Publication No. US20050266420A1 ; GENERAL INFORMATION: ; APPLICANT: PUSZTRI, LAJOS ; APPLICANT: SYMMANS, W. FRASER ; APPLICANT: HESS, KENNETH R. ; APPLICANT: AYERS, MARK ; APPLICANT: STEC, JAMES ; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY ; FILE REFERENCE: UTK:880US ; CURRENT APPLICATION NUMBER: US/10/955, 054A ; CURRENT FILING DATE: 2004-09-30 ; NUMBER OF SEQ ID NOS: 195 ; SOFTWARE: PatentIn Ver. 2.1 ; SEQ ID NO 44 ; LENGTH: 6812 ; TYPE: DNA ; ORGANISM: Homo sapiens US-10-955-054A-44			
 Alignment Scores:			
Pred. No.: 5, 81e-29 Length: 6812			
Score: 359.50 Matches: 121			
Percent Similarity: 38.16 Conservative: 74			
Best Local Similarity: 23.68 Mismatches: 161			
Query Match: 13.46 Indels: 155			
DB: Gaps: 13			
 US-10-618-941-99 (1-507) x US-10-955-054A-44 (1-6812)			
Oy	5	GIUPROAlPRoArxArgAla-----Aarglu 13	
Dd	327	GAGCGGCCCGCCGCGAGAGTGGACGCGCGCGCGAGCGCCAGAGGCCACAGAGAGAA	386
Oy	14	ArggluarGgluarGglunApRgluserGluApRgluserAprileuglugIuserPro 33	
Dd	387	CAGAACCCAGACAGATGATATCGAAAGCTGGAAACCAAGGCCCGTGGAAATGTCTAAC	446
Oy	34	CysglYarGtTgRglInlyeArgrArgglugInValaenGlngIyaenMetProglIyeuGln	53
Dd	447	GATGCGCGCTTCTCAAGTTTGACATCGAAATCGCAGAGGCTCC-----TTTAAG 497	
Oy	54	SerThrPheLeuAlaMeArPrhrGluGluGluValaGluValaITrranGluLeuNh 73	
Dd	498	ACGGTCTCAAAAGTCTGGACACTGAACCAACCCCTGAAAAGTCCCGCTGTGTAAGTCGAC	557
Oy	74	PheglYaPArgrLyueAlaPheAlalaNhlsgIuglInylbegInThrValPheglugIn	93
Dd	558	-----GATCGAAATTTACAAAGTGTGAGAGCCAGCAAGTTTAAADAAGAGTGAATG	611
Oy	94	LeuValleuValaPsrHieProAnlleValLYleuNhileYgTrTrrleuAprhrSer	113
Dd	612	TTAAAGAGCTTCAGCATCCCAATATTTGATGATTATTAAGATTCTGGGAATCCACAGTA	671
Oy	114	GiUAlaCyuaLarYValIllePheIleThrgluTyrValaSerSerGlyeIserleuYvgIn	133
Dd	672	AAAGGAAGAACTGCATTTGTTGGTGAAGCTTAATATAGACGTTGGAAACATTAAAGC	721

OY	134	PheLeuLySvTThrLyLeuAsnHisValAlaSerAlaIAspAlaTrpLyAsnTrp	153
Db	732	TATCTGAAAAGG-----TTTAAAGTATGAAGATCAATGTTCTTAAGACTGG	779
OY	154	CysThrGlnIleLeuSerAlaLeuSerPheLeuHisAlaCysSerProIleIleHis	173
Db	780	TGGCGTCAGATCTTTAAAGCTCTTCAGTTCTTTATCTATCGAATCCACCTATCTTAC	839
OY	174	GlyAsnLeuThrSerAspThrIlePheIleGln--HisAsnGlyLeuIleLySileGly	192
Db	840	CGCGATCTTAAATGTGCACATCTTTATGCCGGCCCTACTGGCTGCTCAAGATTGGA	899
OY	193	SerValTrpHisArgIlePheSerAsnAlaLeuArgProCotHisAlaLeuProAsp	212
Db	900	GACCTC-----GCT	908
OY	213	LeuThrSerProIleArgAlaGlu-----ArgGluGluAspArgAsnHisPhePhe	230
Db	909	CTGGCAACCTGAAAGCGGGCTTTCTTTGSCCAAAGTGTGATGAGTACCACAGATTCTAG	968
OY	231	ProProGluTrpArgGluValAlaAspArgThrAlaValAspIlePheSerPheGlyMet	250
Db	969	GCCCTGAGATGTATGAGAGAAATATGATGAATCCGTGACGTTTAGCTTTGGGATGG	1028
OY	251	CysAlaLeuGluMetAlaValLeuGluIleGlnThrAsnGly-----	264
Db	1029	TGCATGCTTGAGATGGCTACATCTGMAATATCTTACTCGAGTCCCAAAATGCTGGCAG	1088
OY	265	--AspThrArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeuSerAspPro	283
Db	1089	ATCTACCGTCCGCGTACAGTGGGGTGAAGCCAGCCAGCTTTTGACAAAGTGAATCTCT	1148
OY	284	AsnMetArgGluPheIleLeuCysCysLeuAlaArgAspProAlaArgArgProSerAla	303
Db	1149	GAACTGGAAGAAATATTATTGAAGATGCATACGACAAACAAAGATGAAGA-----	1199
OY	304	HisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeuLySLeuLeuAlaAla	323
Db	1200	-----TATTCATCAAAAGACCTTTTGAC	1223
OY	324	HisCysPheIleGlnHisGlnTrpLeuMetProGluAsnValValGluGluLySThrLyS	343
Db	1224	CATGCGTCTTCCAAAGAGAAACAGAGTACGGGTAGAAATTAGCAAGAAAGATGAT---	1280
OY	344	AlaMetAspLeuHisAlaValLeuAlaGluLeuProArgProArgArgProProLeuGln	363
Db	1280	-----	1280
OY	364	TrpArgTrpSerGluValAspPheMetGluLeuAspLySPhenGluAspValArg---	382
Db	1281	-----GGAGAAAATAATGCCATTAATAATTATGGCTACCTATTGAAGATTTAAGAA	1331
OY	383	--AsnGlyIleTrpProLeuMetAsnPheAlaIleThrArgProLeuGlyLeuProArg	401
Db	1332	TTAAAGGGAATAAC-----	1346
OY	402	ValLeuAlaProProGluGluValGlnLyAlaLySThrProThrProGluProPhe	421
Db	1346	-----	1346
OY	422	AspSerGluThrArgLySValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLyS	441
Db	1347	-----AAAGTAAATGAACATATGAGTCTTTTATGATTTAGAGAA-----	1388
OY	442	AlaArgTrpHisLeuThrLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThr	461
Db	1388	-----	1388
OY	462	TyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisTrpGly	481
Db	1389	-----GATGTCCTCCAGAAAGATGTGTCACAAAGAAATGTGATGAGTCTGGG	1430

```
Qy      482 PheLeuH1eG1uAspAaPArpAgtWetLysLeuAla 492
      : : : : : | | | | | : : : : :
Db      1431 TATGTCGTGAAGGTGATCAAGACCAAGCTGCT 1463

RESULT 5
US-11-136-527-688
; Sequence 688, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; PRIOR APPLICATION NUMBER: US 2005-05-25
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 688
; LENGTH: 7555
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-688

Alignment Scores:
Pred. No.:      2,17e-22      Length:      7555
Score:          302.50      Matches:      115
Percent Similarity: 36.26%      Conservative: 71
Best Local Similarity: 22.42%      Mismatches: 170
Query Match:    11.33%      Indels:      157
                        Gaps:      14

US-10-618-941-99 (1-507) x US-11-136-527-688 (1-7555)

Qy      5 GIUProAlaProAArgArg-----AlaArgGluArgGlu 15
      : : : : : | | | | | | | | | | | | | | | | | |
Db      873 GAGCGCCGCCGCCCTCAAGAGTGGAAGTGGCAGCGGTGTCGACAGTCCCAAAAGACCGCGAG 932

Qy      16 ArgGluArgGluAspGluSerGluAsp-----GluSerAAsp1LeuGluGlu 31
      : : : : : | | | | | | | | | | | | | | | | | |
Db      933 GAGGAAGGAGAACCCAGAACAGATGATTCGAAAGAGCTGAGACGACGACAGCATGGGAATG 992

Qy      32 SerProCyGlyArgTyrGlnLysArgGluGluGlnValAlaGlnGlyAspMetProGly 51
      : : : : : | | | | | | | | | | | | | | | | | |
Db      993 TCCAAATGACGCTGCTTTCTCAATTTGACATCGAAATCGCAGAGGCTCC----- 1043

Qy      52 LeuGlnSerThrPheLeuAlaMetAAspThrGluGluGlyValGluValAlaTyrAAspGlu 71
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1044 TTTAAGACGCTGTCAAAAGAGMSKYYRCCGAMWTCACSTGTSAMGMSGCCSTGCTGA 1103

Qy      72 LeuH1ePheGlyAspArgGlyAlaPheAlaAlaH1eGluGluLysLeuGlnThrValPhe 91
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1104 TTGCAG-----GATCGAAAGTTTAAACAAATCTGAAGGACGAGATTTTAAAGAAAGACT 1157

Qy      92 GluGlnLeuValLeuValAspHisProAAsp1LeuValLeuH1eLysValTyrTyrLeuAsp 111
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1158 GAAATGCTAAAGGCTGTCAACATTTTAAACATTTGCTGATGTATCTCGGGAATCC 1217

Qy      112 ThrSerGluAlaCysAlaArgVal1LePhe1LeuThrValSerSerGlySerLeu 131
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1218 ACAAGTAAAGGAGAAAATGCAATGTTTAAAGTGAAGTAAATGACATCTGGAACACTT 1277

Qy      132 LysGlnPheLeuLysLeuThrLysValAspH1eLysValAlaMetAAspAlaArgAlaTyrLys 151
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1278 AAAAGCTACTTAAAGG-----TTTAAAGTATGATAAATCAAGATTTTAAAGA 1325

Qy      152 ArgTyrCyGlyThrGln1LeuSerAlaLeuSerPheLeuH1eAlaCysSerProPro1le 171
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1326 AGCTGGTGTGCGCAATCTTTAAAGGACTTCAATTTCTCATTAACGAACTCCCGGAT 1385

Qy      172 H1eH1eGlyAspLeuThrSerAAspThr1LePhe1LeuGln--H1eAAspGlyLeuLeuLys 190
      : : : : : | | | | | | | | | | | | | | | | | |
```

```
Db      1386 ATTCAACGGGATCTTAAATGTGACAACATCTTTATCACTGGTCTTACCGGCTCAAGTCAAG 1445
Qy      191 H1eGlySerValTyrPhe1Arg1LePheSerAAspAlaLeuArgProProThralaLeuPro 210
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1446 ATCGAGAGAC----- 1454

Qy      211 AspAAspLeuArgSerPro1LeArgAlaGlu-----ArgGluGluLeuArgAAspLeuH1e 228
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1455 YTTGGTCTGGCAACTCTTAAAGCGGCTTTCTTTGCCAAAGTGTATAGTACCCCAAG 1514

Qy      229 PhePheProGluTyrGlyGluValAlaAspGlyThrAlaValAsp1LePheSerPhe 248
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1515 TTTATGGCTCCGAGATGTATGAGAGAAAGTACAGTAAATCCGTTGATGTTATGCTTTT 1574

Qy      249 GlyMetCysAlaLeuGluMetAlaValLeuGlu1LeuGlnThrArgGly----- 264
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1575 GGGATGTCAAGCTTGAAGATGAGTCACTGTAATTCATCTACATCTCAGAGTCCMAATGCT 1634

Qy      265 -----AspThrArgValThrGluGluAla1LeAlaArgAlaArgH1eSerLeuSer 281
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1635 GCTCAGATCTACCGTCCAGTGAAGTCAAGTGAAGAACCCAGCGGTTTGGACAAAGTACGA 1694

Qy      282 AspProAAspMetArgGluPhe1LeuLeuCyGlyLeuAlaArgAspProAlaArgArgPro 301
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1695 ATTCTGAAGTGAAGGAATATTATGAAGAGTATTCGACAAACAAAGATGAAGA--- 1751

Qy      302 SerAlaH1eSerLeuLeuPheH1eArgValLeuPheGluValAlaH1eSerLeuLysLeu 321
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1752 -----TATTCATCAAAAGACTT 1769

Qy      322 AlaAlaH1eCyPhe1LeuGln1LeuGlnTyrLeuMetProGluAAspValValGluGluLys 341
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1770 TTGAACCAATGCCCTTTTCCAGAGAAACAGGGGTACCGGTTGAATTCGACAAAGAAAGAT 1829

Qy      342 ThrLysAlaMetAAspLeuH1eAlaValLeuAlaGluLeuProArgProArgArgProPro 361
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1830 GAT----- 1832

Qy      362 LeuGlnTyrArgTyrSerGluValSerPheMetGluLeuAspLysPheLeuGluAspVal 381
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1833 -----GGAGAAAGATAGCTATCAAAATTTATGGCTAGTATGAAAGATATT 1877

Qy      382 Arg-----AAspGly1LeTyrProLeuMetAAspPheAla1AlaThrArgProLeuGlyLeu 399
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1878 AAGAAATTTAAAGGMAATAC----- 1898

Qy      400 ProArgValLeuAlaProProProGluGluValAlaLysValLysThrProThrProGlu 419
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1898 ----- 1898

Qy      420 ProPheAspSerGluThrArgLysVal1LeuGlnMetGlnCyAAspLeuGluArgSerGlu 439
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1899 -----AAAGCAAGAAAGCTATTGAGTTTCTTTGACCTTGAGAGAG----- 1940

Qy      440 AspLysAlaArgTyrPhe1LeuThrLeuLeuLeuValLeuGluAspArgLeuH1eArgGln 459
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1940 ----- 1940

Qy      460 LeuThrTyrAAspLeuLeuProThrAAspSerAlaGlnAAspLeuAlaSerGluLeuValH1e 479
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1941 -----GATGTACCAAGAAAGTGTTCCTCAAGAAATGCTCGAG 1976

Qy      480 TyrGlyPheLeuH1eGluAspArgArgMetLysLeuAla 492
      : : : : : | | | | | | | | | | | | | | | | | |
Db      1977 TCTGGTATGTCTGTGAAGGTGATCAACAGCAATGCT 2015

RESULT 6
US-10-770-726-27
; Sequence 27, Application US/10770726
; Publication No. US20050266409A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Brown, Eugene
```


Db 156 ----CAGCAGCAGCCGAGAGAAAGAACTCATTATTAATGAGACTCTGGTCAATGAGGAGAAA 211

Qy 98 -AepHisProAsnIleValIlyLeuHnIlybTyTrIrpLeuAspThrSerGIuAlaCybAl 117

Db 212 CAAAAACCCAAACATTGTGAC-----TATCTGACAGATTACTTGTGGGAGA 259

Qy 117 aArgValIlePheIleThGIuTyValSerSerGIySerIleuPheIleuLybLy 137

Db 260 TGAGCTATGCGTGTGTATGGAATACTTGCGCGAGGCTCTCTTAACAGATGTGGTGCAGA 319

Qy 137 sThrIlybLybAsnHnIlyeAlaIeValaIeAsnAlaIaTgAlaIaTrpIlybTyTrpCyrHnGI 157

Db 320 AACCC-----TGATGTGATGAAGGCCAGATAGCAGCTGTGTGCCGAGAGTG 364

Qy 157 eLeuSerAlaLeuSerPheLeuHnIlaIaCybSerProPheIleHnIsgIyAsnLeuTh 177

Db 365 TCTACAAAGCTTTGTGAGTTCCTGCAT-----TCAAACCAAGTCATTTCACAGAGACATCA 418

Qy 177 rSerAspThrIlePheIleGIuHnIsbnGIyLeuIlyeIlybIleGIySerValTrpHisAr 197

Db 419 GAGTGCACATATTCTGCTGGGAATGAGTGTGCTGTCAAAATTAACTGA-CTTGGAGATTCT 477

Qy 197 gIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIl 217

Db 478 GTGCACAGATTACTTC----- 493

Qy 217 eArgAlaGIuArgGIuGIuLeuArg-----AenLeu-HisPhePheProGIuN--T 234

Db 494 -AGAGCAGAGCAGAAAGAGACCATGTGGGAATCTCATTTGGATGGACCTGAAGTTG 552

Qy 234 yTGIyGIuValAlaAspGIyThrAlaValAspIlePheSerPheGIyMetCyaAlaLeuG 254

Db 553 TGACACGCAAGGCGTATGAGCCCAAGATGTGACATCTGTCTCCGGAATTATGGCAATTG 612

Qy 254 lUmeC-----AlaValLeuGIuI 266

Db 613 AATATGTTGAGGGGGAGACCCCATACCTCAATGAAAACTTTGAGAGCCTTGACTCA 672

Qy 260 lGIuHnTrzAnGIyAspThrArgValThrGIuGIuAlaIleAlaArgAlaArgHiserL 280

Db 673 TTGCTAACCATGGG-----ACGCCAGAG-----C 696

Qy 280 eUserAspProAsn-----MetArgGIuPheIleLeuCybCysLeuAlaA 295

Db 697 TTCAGAACCCAGAGAGATTGTCACTATTTCCGGGACCTTTGAAACCGCTGTCTTGAGA 756

Qy 295 rGAspProAlaArgArgProSerAlaHisSerIleuPheHnIsbnArgValLeu 312

Db 757 TGGATGTGAGAGAGAGAGCTTCAGCTTAAGAGACTGACAGCATCAATTCCTG 809

RESULT 10

US-11-134-563-15

Sequence 15, Application US//11134563

Publication No. US20050287569A1

GENERAL INFORMATION:

APPLICANT: Leong, John M.

APPLICANT: Campellone, Kenneth G.

TITLE OF INVENTION: ESPRU NUCLEIC ACIDS AND PROTEINS AND

TITLE OF INVENTION: USES THEREOF

FILE REFERENCE: 07917-280001

CURRENT APPLICATION NUMBER: US/11/134,563

CURRENT FILING DATE: 2005-05-20

PRIOR APPLICATION NUMBER: US 60/573,600

PRIOR FILING DATE: 2004-05-20

NUMBER OF SEQ ID NOS: 26

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 15

LENGTH: 1638

TYPE: DNA

ORGANISM: Homo sapiens

US-11-134-563-15

Alignment Scores:

[illegible]


```

: TYPE: DNA
: ORGANISM: Homo sapiens
US-10-770-726-39

Alignment Scores:
Pred. No.: 1,88e-08 Length: 2253
Score: 173.00 Matches: 86
Percent Similarity: 40.23% Conservative: 54
Best Local Similarity: 24.71% Mismatches: 131
Query Match: 6.48% Indels: 78
DB: 6 Gaps: 16

US-10-618-941-99 (1-507) x US-10-770-726-39 (1-2253)

QY 7 AAlPcRArgArAlaArgLuArgLuArgLuArgLuSerGluSerGluSer 26
   |||||
Db 450 TCTCCAGTCACAGCCGTTTCAGAAATCAGAGCAGAAATTCGACGACCACTGATC 509
   |||||
QY 27 AspIleLeuGlu-----GluSerProCysGlyArg 36
   |||||
Db 510 CTCATTCCTCTCTCCAGGCCACTGAATTAACCCAAAAGACAGCAAGCCCTGCTGCACTCG 569
   |||||
QY 37 -----TrpGlnLysArgArgGluGlnValAsnGlnGly 47
   |||||
Db 570 CACCTGAAAATATATCCTGAGAGAACTGGCATCAAAACAGAAAATGAAATCAAAA 629
   |||||
QY 48 AsnMetProGlyLeuGlnSerThr-----PheLeuAlaMetAspThrGluGluGly--- 64
   |||||
Db 630 AGAGGCACTGGGCTTTGGAAGACTTTGAATTTGTCGCCCTCGGTAAAGAAAGTTTG 689
   |||||
QY 65 ValGluValValTrpAsnGlu----- 71
   |||||
Db 690 GTAAATGTTATTGGCAAGAAAGCAAAAGCAAGTTATTCTGCGCTTTAAAGTGTAT 749
   |||||
QY 72 LeuHisPheGlyAspArgLysAlaPheAlaAlaHis-----GluGluLysIle 87
   |||||
Db 750 TTTAAAGCTCAGCTG--GAGAAAGCCGAGTGGAGCATCAGCTCAGAAAGAGATGAAATA 808
   |||||
QY 88 GlnThrValPheGluGlnLeuValLeuValAspHisProAsnIleValLysLeuHisLys 107
   |||||
Db 809 CAGTCCCACTTCGG-----CATCTTAATATTCTTAGACTGTATGCT 850
   |||||
QY 108 TyrTrpLeuAspThrSerGluAlaCysAlaArgValIlePheIleThrGluTyrValSer 127
   |||||
Db 851 TATTTCCATGATGCTAAC-----AAGAGTACCTTAATCTGGAATATGCACCA 898
   |||||
QY 128 SerGlySerLeuLysGlnPheLeuLysLysThrLysLysAsnHisLysAlaMetAsnAla 147
   |||||
Db 899 CTTGGAAACAGTTTATGAGAACTTCGAAACTTTGAAAGTTTGATGAGCAGAGAACTGCT 958
   |||||
QY 148 ArgAlaIleTrpLysArgTrpCysThrGlnIleLeuSerAlaLeuSerPheLeuHisAlaCys 167
   |||||
Db 959 -----ACTTATATAACAGAAATGCGCAAAATGCCCTGCTTACTGTCAAT----- 1000
   |||||
QY 168 SerProIleIleHisGlnLysAsnLeuThrSerAspThrIlePheIleGlnHisAsnGly 187
   |||||
Db 1001 TCGAAGAGAGTTTATTCATAGACATTAAGCAAGAACTTACTTTGGATCAGCTGGA 1060
   |||||
QY 188 LeuIleLysIleGlySerValTrpHisArgIlePheSerAsnAlaLeuArgProProThr 207
   |||||
Db 1061 GAGCTTAAATTTGCAAT-----TTTGGGTGTCAGTACATGCTCCTCATCT 1105
   |||||
QY 208 AlaLeuProAspAspLeuArgSerProIleArgAlaGluArgGluGlnLeuArgAsnLeu 227
   |||||
Db 1106 TCCAGAGAGACCACTCTC-----TGTGGCACCCCTG 1135
   |||||
QY 228 HisPhePheProProGluLys-----GlyGluValAlaAspGlyThrAlaValAspIle 245
   |||||
Db 1136 GACTACCTGCCCCCTGAAATGATGATGAGGTGGATGCATGAT--GAGAAAGGTGATCTTC 1192
   |||||
QY 246 PheSerPheGlyMetCysAlaLeuGluMetAlaValLeuGluIleGlnThrAsnGlyAsp 265
   |||||
Db 1193 TGGAGCCTTGAAGTCTTGTTCATATGAAATTTTAAGTTGGAGACCTCTTTGAGGCAAC 1252

```

```

QY 266 Thr---ArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeuSerAsp----- 282
   |||||
Db 1253 ACATACCAAGAGACCTTACAAAAGAAATATCACGGCTTGAATTCACATTCCTGACTTTGTA 1312
   |||||
QY 283 ---ProAsnMetArgGluPheIleLeuCysCysLeuAlaArgAspProAlaArgArgPro 301
   |||||
Db 1313 ACAAGAGGAGCCAGGAGACTTCTATTCAAGACTGTTGAAGCATTAATCCACGACGAGGCCA 1372
   |||||
QY 302 SerAlaHisSerLeuLeuPheHis 309
   |||||
Db 1373 ATGCTCAGAGAAAGTACTTGAACAC 1396
   |||||

```

Search completed: January 14, 2006, 03:09:55
 Job time : 407 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 13, 2006, 23:36:39 ; Search time 965 Seconds

(without alignments)
4344.637 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670
Sequence: 1 MAAPPPARRRRERERERED.....RMKLAFLSTPLKTRGTA 507

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-Model=frame+ p2n.model -Dev=xlp
-Q=/cgn2_1/USPTO_spool_p/US10618941/runat_12012006_070206_14384/app_query.fasta_1.647
-DB=Published Applications NA Main -QFMT=Fastap -SUFFIX=rnpbm -MINMATCH=0.1
-LOOCL=0 -LOOEXT=0 -UNITS=bits -STRAT=1 -END=-1 -MATRIX=blowsum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR_SCORE=spec -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pio -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10618941@cgn_1.1.1549@runat_12012006_070206_14384
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA Main:

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2670	100.0	3765	8	US-10-618-941-33
2	2584	96.8	3837	7	US-10-425-114-26870
3	2584	96.8	3838	7	US-10-425-114-26848
4	2584	96.8	3846	7	US-10-425-114-26873
5	2531.5	94.8	3128	7	US-10-425-114-28974
6	2499	93.6	3147	9	US-10-840-512-71
7	2368	88.7	3538	5	US-10-175-523-117
8	2368	88.7	3538	10	US-11-099-266-117

9	1617	60.6	981	3	US-09-862-027-9	Sequence 9, Appl
10	1617	60.6	981	3	US-10-989-228-9	Sequence 9, Appl
11	1534.5	57.5	2029	3	US-09-840-787-69	Sequence 69, Appl
12	1534.5	57.5	2163	5	US-10-153-668-321	Sequence 321, Appl
13	1534.5	57.5	2163	9	US-10-887-553A-882	Sequence 882, Appl
14	1534.5	57.5	2163	9	US-10-956-157-1845	Sequence 1845, Ap
15	1534.5	57.5	2225	5	US-09-764-878-151	Sequence 151, App
16	1529.5	57.3	2159	5	US-10-059-588-1	Sequence 1, Appl
17	1522.5	57.0	2178	5	US-10-153-668-433	Sequence 433, App
18	1508.5	56.5	2241	5	US-10-040-647-9	Sequence 9, Appl
19	1502.5	56.3	1931	3	US-09-925-297-150	Sequence 150, App
20	1174.5	44.0	2803	10	US-11-097-143-269	Sequence 269, App
21	1022.5	38.3	5655	10	US-11-097-143-268	Sequence 268, App
22	955	35.8	1065	3	US-09-764-868-185	Sequence 185, App
23	955	35.8	1065	3	US-09-764-875-455	Sequence 455, App
24	807.5	30.2	1400	9	US-10-956-157-7080	Sequence 7080, App
25	786.5	29.5	1400	10	US-11-060-756-140	Sequence 140, App
26	786.5	29.5	1400	10	US-11-060-756-141	Sequence 141, App
27	786.5	29.5	1400	10	US-11-060-756-4412	Sequence 4412, App
28	786.5	29.5	1400	10	US-11-060-756-4413	Sequence 4413, Ap
29	756	28.3	538	6	US-10-006-285-456	Sequence 456, App
30	756	28.3	910	3	US-09-764-868-209	Sequence 152, App
31	756	28.3	910	3	US-09-764-875-152	Sequence 152, App
32	669	25.1	2105	6	US-10-006-285-280	Sequence 280, App
33	651	24.4	862	3	US-09-814-353-20005	Sequence 20005, A
34	490.5	18.4	2546	6	US-10-104-047-1939	Sequence 1939, Ap
35	424.5	15.9	509	3	US-09-814-353-14436	Sequence 14436, A
36	423	15.8	477	3	US-09-814-353-1699	Sequence 1699, Ap
37	423	15.8	477	3	US-09-814-353-8052	Sequence 8052, Ap
38	413.5	15.5	19533	6	US-10-292-798-1113	Sequence 1713, Ap
39	412.5	15.4	13630	3	US-09-764-868-1369	Sequence 1369, Ap
40	412.5	15.4	13630	3	US-09-764-875-1220	Sequence 1220, Ap
41	412.5	15.4	21221	6	US-10-017-161-2067	Sequence 2067, Ap
42	397.5	14.9	2404	6	US-10-353-929-58	Sequence 58, Appl
43	397.5	14.9	4415	7	US-10-433-794-37	Sequence 37, Appl
44	397.5	14.9	6582	9	US-10-490-597-6	Sequence 6, Appl
45	397.5	14.9	7825	9	US-10-490-592-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-618-941-33
; Sequence 33, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 33
; LENGTH: 3765
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-618-941-33

Alignment Scores:
Pred. No.: 2,996-313
Score: 2670.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 8
US-10-618-941-99 (1-507) x US-10-618-941-33 (1-3765)

Length:	3765
Matches:	507
Conservative:	0
Mismatch:	0
Indels:	0
Gaps:	0

```

Oy      1 MetAlaAlaProGluProAlaProArGrAlaArgGluArgGluArgGluAsp 20
Db      282 ATGGGAGCCCGGAGACCAGGAGCGGCGCGGGAACGGAGCGGAGGAGGAGAC 341
Oy      21 GluSerGluAerGluSerAsp11eLuglUgluSerProCyGslYArgTrpGlnLysArg 40
Db      342 GAGAGCGAGGACGAGAGCGACATCTTGAGGAAAGCCCGTGGTGGTGGTGGCAAAAGCGA 401
Oy      41 ArgGluGlnValaerGlnGlyAsmMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
Db      402 CGGAGCGAGGTAAACCAAGGAAACATGCCAGGCTTCAGAGACATCTTCCTGACCATGAC 461
Oy      61 ThrGluGluGlyValaGluVala1TrpAsnGluLeuHisPheGlyAspArgLysAlaPhe 80
Db      462 ACGAGAGGAGGGGTGAGAGTGGTGTGAAACAGCTCCATTCGAGAGACAGAGAGCCCTTC 521
Oy      81 AlaAlaHisGluGluGlyVala1eGlnThrValaPheGluGlnLeuValaLeuValaAspHisPro 100
Db      522 GCGGCGCAGAGAGAAATCCAGACCGGTTCAGAGACGTGGTGGTGGTGGTGGTGGTGGTGG 581
Oy      101 Asn11eValaLysLeuHisLysGlyTrpLeuAspThrSerGluAlaCyAlaArgVala11e 120
Db      582 AACATCGTGAAGTTCACAAAGTACAGCTGATACCTCGAGGCTTCGCGCAGAGGTCTATC 641
Oy      121 Phe11eThrGluGlyValaSerSerGlySerLeuLysGlnPheLeuLysValaSerThrLys 140
Db      642 TTTCATCAGAGTACGTATCATCAGGACGCTCAAGCAATTCCTCAAAAGACCAAGAG 701
Oy      141 AsnHisLysValaMetAsnAlaArgAlaTrpLysArgTrpCySerThrGln11eLysSerAla 160
Db      702 AACCAAGAGGCGATGAGACCGCGGCGCTGAAAGCGCTGGTGGACGACATCTGCTGCG 761
Oy      161 LeuSerPheLeuHisAlaCySerProPro11eHisGlyAsnLeuThrSerAspThr 180
Db      762 CTCAGCTTCCTCGACAGCTCGACGCCCAATCATCCAGGAACTCAACACGACGACACC 821
Oy      181 IlePhe11eGlnHisAsnGlyLeuLeu11eLys11eGlySerVala1TrpHisArg11ePheSer 200
Db      822 ATCTTCATTCAGCACAGGAGCGCTCATCAAGATCGCTCGTGGTGGACGAGATCTTCTCC 881
Oy      201 AsnAlaLeuArgProProThrAlaLeuProAspAspLeuAspSerPro11eArgAlaGlu 220
Db      882 AATGCTCGCGGCGCTCCACAGACATTCAGATGATCCGAGAGCCCATCGGCGCTAG 941
Oy      221 ArgGluGluLeuAspAsnLeuHisPhePheProProGlyTrpGlyVala11aAspGly 240
Db      942 CGAGGAGAACTTCGAACTCGACTTCTCCCGGAGATGAGAGTGGAGTGGCGCATGG 1001
Oy      241 ThrAlaValaAsp11ePheSerPheGlyMetCysAlaLeuGluMetAlaValaLeuGluTrp 260
Db      1002 ACCGCTGTGACATTTCTCTCTTGGATGTGTGGCTGGAGATGGCTGTACTGGAATTC 1061
Oy      261 GlnThrAsnGlyAspThrArgVala1ThrGluGluAla11eAlaArgAlaArgHisSerLeu 280
Db      1062 CAGACCAATGGGAGACACCGGAGTCAAGAGAGGCGCATTTGCTCGGACGACGACTCGCTG 1121
Oy      281 SerAspProAsnMetArgLuphe11eLeuCyGlyLeuAlaArgAspProAlaArgArg 300
Db      1122 AGTGACCCCAACATGCGGAGATTCATCTTGTGCTGCTGGCGCGGAGACCTCGCGCGCG 1181
Oy      301 ProSerAlaHisSerLeuLeuPheHisArgVala1LeuPheGluValaHisSerLeuLysLeu 320
Db      1182 CCTCTGCGCACAGCTCTCTCTTCCACGCGGTCTTTCGAGGTGACTCGTGAAGTTC 1241
Oy      321 LeuAlaAlaHisCyPhe11eGlnHisGlnTrpLeuMetProGluAsnVala1GluGlu 340
Db      1242 CTGGAGGCGCACTGCTTCATCCAGACCAAGTACCTCAAGCTCGAGAACTGTGTGAGAGAG 1301
Oy      341 LysThrLysValaMetAspLeuHisAlaValaLeuAlaGluLeuProArgProArgArgPro 360
Db      1302 AAGACCAAGGCGCATGAGACTGACCGGCTGTGGCGAGACTTCCCGGCGCGCAGGCGCC 1361
Oy      361 ProLeuGlnTrpArgTrpSerGluVala1SerPheMetGluLeuAspLysPheLeuGluAsp 380

```

```

Db      1362 CCGCTGACGTGGCCGCTACTCGGAAGTCTCTTCATGAGCTGAGCAAAATTCCTGGAGAT 1421
Oy      381 ValArgAsnGly11eTrpProLeuMetAsnPheAla1aThrArgProLeuGlyLeuPro 400
Db      1422 GTCAAGAAATGGAATCTAACCTAGTATGAATTTTCACGACACTGACCCCTGGGCTGCCC 1481
Oy      401 ArgValaLeuAlaProProProGluGluVala1GlnLysAlaLysThrProThrProGluPro 420
Db      1482 CGTGTGCGCCCAACCCCGAGAGAGTCCAAAGGCGAAAGACCCGACCCAGAGGCC 1541
Oy      421 PheAspSerGluThrArgLysVala11eGlnMetGlnCysAsnLeuGluArgSerGluAsp 440
Db      1542 TTTCATCTGAGACCAAGAGTATCATCAGATGACAGTCCAACTCGAGAGAAAGAGAGAC 1601
Oy      441 LysAlaArgTrpHisLeuThrLeuLeuVala1eGluAspArgLeuHisArgGlnLeu 460
Db      1602 AAGGCGCGTGGCATCTCACTGCTGCTGGTGGAAAGACCGGCTGACCGGACGCTG 1661
Oy      461 ThrTrpAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValaHisTrp 480
Db      1662 ACCTAGCACTGTCTCCCAACGAGACGAGCGCCAGGACCTCGCTGGAGCTGTGCATAT 1721
Oy      481 GlyPheLeuHisGluAspAspArgMetLysLeuAla1aPheLeuGluSerThrPheLeu 500
Db      1722 GCTTCTCTCCACGAGAGACGACCGGATGAGCTGGCCGCTTCTCGAGAGACCTTCTC 1781
Oy      501 LysTrpArgGlyThrGlnAla 507
Db      1782 AAGTACCGTGGAGCCAGGCC 1802

RESULT 2
US-10-425-114-26870
; Sequence 26870, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovacic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26870
; LENGTH: 3837
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4656-003-G12_FLI
US-10-425-114-26870

Alignment Scores:
Pred. No.: 9.2e-303 Length: 3837
Score: 2584.00 Matches: 501
Percent Similarity: 88.05% Conservative: 0
Best Local Similarity: 88.05% Mismatches: 0
Query Match: 96.78% Indels: 68
DB: Gaps: 2

US-10-618-941-99 (1-507) x US-10-425-114-26870 (1-3837)
Oy      1 MetAlaAlaProGluProAlaProArGrAlaArgGluArgGluArgGluAsp 20
Db      231 ATGGGAGCCCGGAGACCAGGAGCGGCGCGGGAACGGAGCGGAGGAGGAGAC 290
Oy      21 GluSerGluAerGluSerAsp11eLuglUgluSerProCyGslYArgTrpGlnLysArg 40
Db      291 GAGAGCGAGGACGAGAGCGACATCTTGAGGAAAGCCCGTGGTGGTGGTGGTGGTGGTGG 350

```

```

QY 41 ArgGluGlnValAsnGlnGlnValAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
DB 351 CGGAGCAAGTAAACCAAGGAAACATGCAAGGCTTCAAGACACTTCTAGCAATGAGAC 410
QY 61 ThrGluGluGlyValGluValValTPAangluLeuHisPheGlyAspArgLysAlaPhe 80
DB 411 ACGGAGAGAGGGGTAGAGGTGTGTGTGAACGAGCTCACACTTGGAGACAGAGAGGCTTC 470
QY 81 AlaAlaHisGluGluGluValIleGlnThrValPheGluGlnLeuValIleAspPro 100
DB 471 GCGGGGCAAGAGAGAAAGATCCAGACCGGTGTGAGACAGCTGGTCTGGTGAACACCGG 530
QY 101 AsnIleValIleLeuHisValSerTyrTrpLeuAspThrSerGluAlaCysAlaArgValIle 120
DB 531 AACATCGTAAATTGACAAAGTACGTGGTGAATACCTCTGAGGCTCGCGAGGCTCATC 590
QY 121 PheIleThrGluTyrValSerSerGlySerLeuLysGlnPheLeuLysIleThrLysLys 140
DB 591 TTCATCAGAGATGAGTGTGCATCAGGCAAGCTCAAGCAATTCCTCAAAAGACAAAGAG 650
QY 141 AsnHisLysValMetCysAsnAlaArg----- 148
DB 651 AACCAAGGCGCATGAACCGCGGTATGGGAGCGGCTGGGAGACCAAGGAGACAG 710
QY 148 ----- 148
DB 711 ACGGGGTGGGAGACCTCGGGAGCTGGAGTGTGAGGGGTGCGCGGCGCTCGGACA 770
QY 148 ----- 148
DB 771 GGGGCTGGGCGAGATGCGGGGCGGCTCGGAGCGGCGGCTCTCTCGGCCACAC 830
QY 149 ----- 148
DB 831 CGAGCCAGCGAGTCTGCTGCTCGCGCGGCGGCTGGAGCGCTGGTGCAGCAGATCTG 890
QY 159 SerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeuThrSer 178
DB 891 TCTGGCTCAGCTTCTCTGCAAGCTGCGACCGCCCAATCATCACGGGAACTGACCAAC 950
QY 179 AspThrIlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValTrpHisArgIle 198
DB 951 GACACCATTTTATTCAGACAAACGGCTCATCAAGATGGCTCGTGGGACCGAATC 1010
QY 199 PheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArg 218
DB 1011 TTCTCCAAAT-----GCACTTCCAGATGATCTCGAAGCCCATTCGC 1052
QY 219 AlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGluValAla 238
DB 1053 GCTGAGCAGAGAACTTCGAACTGCACTTCTTCCCGCCAGATATGAGAGGTGGCC 1112
QY 239 AspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeu 258
DB 1113 GATGGAGCCGCTGTGACATCTTCTCTTGGAGTGTGTGGCCCTGAGATGGCTGTACTG 1172
QY 259 GluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHis 278
DB 1173 GAATCCAGAACCAATGGGAGACACCGGGTCAAGAGAGGCCATTGGCTGGCCAGGCAC 1232
QY 279 SerLeuSerAspProAsnMetCysGluPheIleLeuCysValLeuAlaArgAspProAla 298
DB 1233 TCGGAGATGACCCCAACATGGGAGATTCATCTTGTGTGCTGGCCCGGAGCCCTGCG 1292
QY 299 ArgArgProSerAlaHisSerLeuPheHisArgValIlePheGluValHisSerLeu 318
DB 1293 CGCGGCGCTCTGCGCCACAGCTCTTCCACCGCGTGTCTTGAAGTGCATCTGCGG 1352
QY 319 LysLeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetProGluAsnValVal 338
DB 1353 AAGCTCTGGAGCCCACTGCTTATCCAGACCAAGTACCTGCTGAGAAATGTGGTG 1412

```

```

QY 339 GluGluValThrLysAlaMetAspLeuHisAlaValLeuAlaGluLeuProArgProArg 358
DB 1413 GAGAGAAAGACCAAGCCATGAGACCTGACAGCGGTCTTGGAGAGCTTCCGCGCCCGC 1472
QY 359 ArgProProLeuGlnThrArgTyrSerGluValSerPheMetGluLeuAspLysPheLeu 378
DB 1473 AGGCCCCCGCTCAGTGGGCGGTACTCGAAGTCTCTTCAATGAGCTGACAAATTCCTG 1532
QY 379 GluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAlaThrArgProLeuGly 398
DB 1533 GAGATGTGAGAAATGAAATCTAACCACTGAAGCACTTTCAGACCACTGAGACCTGGGG 1592
QY 399 LeuProArgValLeuAlaProProProGluGluValGlnLysAlaLysThrProThrPro 418
DB 1593 CTGCCCCGTGTCTGGCCCCACCGCGAGAGGTCCAAAGGCCCAAGACCCCGAGCGCA 1652
QY 419 GluProPheAspSerGlyThrArgLysValIleGlnMetGlnCysAsnLeuGluAspSer 438
DB 1653 GAGCCCTTGACTGTGAGACCAAGAAAGTCAATCCAGATCCAGATGCAACTGGAGAGAGC 1712
QY 439 GluAspLysAlaArgTyrPheIleLeuThrLeuLeuValLeuGluAspArgLeuHisArg 458
DB 1713 GAGGACAAAGCGCGCTGGCATCTCACTGTCTGTGTGCTGGAAGACCGGCTGCACCG 1772
QY 459 GlnLeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuVal 478
DB 1773 CAGCTGACCTTACGACCTGTCTCCAAAGCAGCGGCCAGAGACTCGCTCGAGCTCGG 1832
QY 479 HisTyrGlyPheLeuHisGluAspAspArgMetLysLeuAlaAlaPheLeuGluSerThr 498
DB 1833 CACTATGCTTCTCTCCACAGAGACGAGCGGATGAAAGTGTGGCCGCTTCTGAGAGACCT 1892
QY 499 PheLeuLysTyrArgGlyThrGlnAla 507
DB 1893 TTCCTCAAGTACCGTGGAGCCAGGCC 1919

```

RESULT 3

```

US-10-425-114-26848
; Sequence 26848, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jindong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26848
; LENGTH: 3838
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4654-015-P10_FLI
US-10-425-114-26848

```

Alignment Scores:

```

Pred. No.: 9,2e-303 Length: 3838
Score: 2584.00 Matches: 501
Percent Similarity: 88.05% Conservative: 0
Best Local Similarity: 88.05% Mismatches: 0
Query Match: 96.78% Indels: 68
DB: 7 Gaps: 2

```

US-10-618-941-99 (1-507) x US-10-425-114-26848 (1-3838)

```

QY 1 MetAlaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluAsp 20
|||||

```


Score: 2531.50 Matches: 499
Percent Similarity: 83.47% Conservative: 1
Best Local Similarity: 83.31% Mismatches: 1
Query Match: 94.81% Indels: 99
DB: 7 Gaps: 3
US-10-618-941-99 (1-507) x US-10-425-114-28974 (1-3128)

QY 1 MetAlaAlaProGluProAlaProArgArgAlaArgGluArgGluArgGluArg 20
Db 120 ATGGGGGCCCCGGAGCCGGCCGAGGGGGCCGGGAACGGAGCCGGAGCCGGAGGAC 179
QY 21 GluSerGluAapGluSerAapIleuGluGluSerProCySGlyArgTrpGlnlyArg 40
Db 180 GAGAGCGAGAGAGAGGAGCATCTCGAGAGAAAGCCCTGTGTGCTGGCAAAAGCGA 239
QY 41 ArgGluGluValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAap 60
Db 240 CGGGAGCAGTAAACCAAGGGAACATGCCAGGGCTTCAGAGCACTTCCTAGCCATGGAC 299
QY 61 ThrGluGluGlyValGluValValTTPaNGluLeuHisPheGlyAapArgGlyAlaPhe 80
Db 300 ACGGAGGAGGGGGTAAGAGTGGTGGAAACAGCTCCACTTCGAGAGACGAGAGCCCTTC 359
QY 81 AlaAlaHisGluGluValIleGlnThrValPheGluGlnLeuValLeuValAapHisAap 100
Db 360 GCGGGCGACGAGAGAAATCCAGACCTGTTCAGACAGCTGGTGTCTGTGGACCAACCG 419
QY 101 AsnIleValIleuHisIleValTTPleuAapThrSerGluAlaCyAlaArgValIle 120
Db 420 AACATCGGAAGTTCACAAAGTACGTGCTGATACCTCTGAGGCGCTGGCGAGGCTATC 479
QY 121 PheIleThrGluTyrValSerSerGlySerLeuGlnPheLeuValIleValIleVal 140
Db 480 TTCATCACAAGTACGTGTCTCATCAGGCAAGCTCAAGCAATTCCTCAAAAGACCAAGAG 539
QY 141 AsnHisIleValMetAapAlaArg----- 148
Db 540 AACCAACAGGCCATGAACGCCCGGGTATGGGAGCGGGCTGGGGACGCCACGGGAGACGG 599
QY 148 ----- 148
Db 600 ACGGGGTGGGGCAGACTCGGGACCTGGGATGTGAGGGGGTGGCCCGGCGCTCGAGCA 659
QY 148 ----- 148
Db 660 GGGGCTGGGCGAGATGCGGGGCGGGCTCCGACAGCCAGCCGCTCTCGCGCCACCC 719
QY 149 -----AlaTTPlyAapTrpCyAThrGlnIleLeu 158
Db 720 CGACCGAGGAGTGTGCTCGCGCCGACAGGCTTGGAAAGCGCTGGTGCACCAAGTCTTG 779
QY 159 SerAlaLeuSerPheLeuHisAlaCySerSerProIleIleHisGlyAsnLeuThrSer 178
Db 780 TCTGCGCTCAGACTTCTCTCAGCGCTGAGCCGCCCAATCATCCACGGGAACTTGACGACG 839
QY 179 AspThrIlePheIleGlnHisAsnGlyLeuIleValIleGlySerValTTPHisArgIle 198
Db 840 GACACCATCTTATTCAGACCAACCGCTCATCAAGATCGGCTCGGTGTGCAACGAAATC 899
QY 199 PheSerAapAlaLeuArgProProThrAlaLeuProAapAapLeuArgSerProIleArg 218
Db 900 TTCTCCAAAT-----GCACTTCCAGATGATCTCCGAAGCCCAATCCGC 941
QY 219 AlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGluValAla 238
Db 942 GCTGAGCAGAGAGAACTTCGGAACCTGCACCTTCTCCGCCAGAGTAAAGAGGTGGCC 1001
QY 239 AapGlyThrAlaValAapIlePheSerPheGlyMetCyAlaLeuGluMetAlaValLeu 258
Db 1002 GATGGAGCCGCTGGAGCATCTTCTCTTTGGAGATGTGCGCTGGAGATGCTGTACTG 1061
QY 259 GluIleGlnThrAanGlyAapThrArgValThrGluGluAlaIleAlaArgAlaArgHis 278

Db 1062 GAATTCAGACCAATGGGGAGACCCGGGTCAAGAGAGGCCATTTGCTCGGCCAGGAC 1121
QY 279 SerLeuSerAapProAanMetArgGluPheIleLeuCyValLeuAlaArgAapProAla 298
Db 1122 TCGCTGAGTACCCCAACATGCGGGAGTTCATCTTTGCTGTGCTGGCCCGGAGCCCTGCC 1181
QY 299 ArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeu 318
Db 1182 CGCGGCGCTTGGCCACAGCTCTCTTCCACCGCGGTCTTCGAGGTGCACTGCTG 1241
QY 319 LysLeuLeuAlaAlaHisCyPheIleGlnHisGlnTyrLeuMetProGluAanValVal 338
Db 1242 AAGCTCTGGCAGCCCACTGCTTATCAGACCAACAGTACCTCAAGCTCGAAGAAATGTG 1301
QY 339 GluGluValThrIlyAlaMetAapLeuHisAlaValLeuAlaGluLeuProArgProArg 358
Db 1302 GAGGAGAAAGCAAGGCCATGACCTGACCGCGGTCTTGGCGGAGCTTCCCGGCCCGGC 1361
QY 359 ArgProProLeuGlnTTPArgTyrSerGluValSerPheMetGluAapArgPheLeu 378
Db 1362 AGGCCCCCGTCAGTGGCGGTACTGCAAGTCTCTTCATGAGGTGAGCAAAATTCCTG 1421
QY 379 GluAapValArgAanGlyIleTyrProLeuMetAapPheAlaAlaThrArgProLeuGly 398
Db 1422 GAGGATTCAGAAATGGAATCTACCACTGATGAATTTGACAGCACTCGACCCCTGGGG 1481
QY 399 LeuProArgValLeuAlaProProProGluGluValGlnIlyValIlyThrProThrPro 418
Db 1482 CTGCCCCGTGTGCTGGCCCCACCCCGAGAGAGGTCCAAAGGCCAAGACCCGACGCCA 1541
QY 419 GluProPheAapSerGluThrArgGlyValIleGlnMetGlnCyValLeuGluAapSer 438
Db 1542 GAGCCTTTGACTCTTGAGACCAAGAAAGTCAATCAGATGACAGTGCACCTGGAGAGAGC 1601
QY 439 GluAapValAaArgTTPHisIleuThrLeuLeuValIleuGluAapArgLeuHisArg 458
Db 1602 GAGGACAAAGCGCGCTGACATCTCATCTGCTGTGCTGGAAGACCGGCTGACCGG 1661
QY 459 GlnLeuThrTyrAapLeuLeuProThrAapSerAlaGlnAapLeuAlaSerGluLeuVal 478
Db 1662 CAGCTGACCTACGACCTGCTCCCAACGAGACGCGCCAGAGACTCGCTCGGAGCTGCTG 1721
QY 479 HisTyrGlyPheLeuHisGlu----- 485
Db 1722 CACTATGCTTCTTCACGA-GGTGCGTGGCGGTGCGCGGCTTGCGGAGGGGCGC 1780
QY 486 -----AapAapArg 488
Db 1781 GCACGGGGCAGCGCGCCCTCCGTCCCAATGCTCCCTCTCCGCGAGAGACCGG 1840
QY 489 MetIleLeuAlaAlaPheLeuGluSerThrPheLeuValTyrArgGlyThrGlnAla 507
Db 1841 ATGAAGCTGCGCGCTTCTGTGAGAGCACTTCTCAAGTACGTTGGAGAACCCAGGCC 1897

RESULT 6
US-10-840-512-71
; Sequence 71, Application US/10840512
; Publication No. US20050125852A1
; GENERAL INFORMATION:
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: MANNING, GERARD
; APPLICANT: CHARIDCZAK, GLEN
; APPLICANT: GRIGORIEV, IGOR
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-1455
; CURRENT APPLICATION NUMBER: US/10/840,512
; CURRENT FILING DATE: 2004-05-07
; PRIOR APPLICATION NUMBER: 60/469,014
; PRIOR FILING DATE: 2003-05-09
; NUMBER OF SEQ ID NOS: 239
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 71

Db	950	AGTACACCCCAATGACGGGAATTATCATCTCTCTGCGCCGACCGGAACTTCGCCCGCA	1008
Qy	301	ProSer1aAh1sSerLeuLeuPheH1sArgValLeuPheGluVal1HsSerLeuLysLeu	320
Db	1010	CCCTCAGCCCAACCTCTCTCTTCCACCGAAGTGCTCTTGGAGGAGCACTCTCTAAAGCTG	1068
Qy	321	Leu1a1aH1sCysPhe1aG1nH1sG1nTy1rLeuMeCProG1uAsnVal1aG1uG1u	340
Db	1070	CTGGCAGCTCATCTGCTTCATCTCAGACCAAGTAACTCATGCTTGAGAAATGTGTAGAGAA	1128
Qy	341	LysThr1yrsalameAspLeuH1sAlaValLeu1aGluLeuProArgProArgArgPro	360
Db	1130	AAGACCAAGCCATGAGACCTCCATGACGATTTTGGCTGAGANTGCCGACCCCATGAGACC	1168
Qy	361	ProLeuG1nT1rPa1gTy1rSerG1uVal1SerPheMeC1uLeuAspLysPheLeuG1uAsp	380
Db	1190	CCAAATGCAAGTGGCCGTA1CTCAGAGAGTCTCTTCTTGGAGCTGAGCAAAATTCCTAGAGAT	1248
Qy	381	Val1ArgAsnG1y1LeTy1rProLeuMeCAsnPhe1a1aThrArgProLeuG1uLeuPro	400
Db	1250	GTCAGAAACGGGATCTATTCACATGATGAATCTTGCGCTGCTGCGCCCTTGGGCGTTCC	1308
Qy	401	ArgVal1leu1aP1rProProG1uG1uVal1G1nLys1aLys1rProTh1rProG1uPro	420
Db	1310	CGTGTGTTGGCCCAACCCCAAGAGAAACCCAAAGGCCAAAATCTCCACGCCCAAGACC	1368
Qy	421	PheAspSerG1uThr1rArgLysVal11eG1nMeC1nCy1sAsnLeuG1uArgSerG1uAsp	440
Db	1370	TTTGACTCGAGAGACCAAGAAAGTGATGTCAGATGACATGCAACTGGAAAGAGCGAGAC	1428
Qy	441	Lys1a1ArgTrpH1sLeuThr1rLeuLeuLeuVal1leuG1uAspArgLeuH1sArg1nLeu	460
Db	1430	AAGGCTGGTGAGCACTTACTCTGCTCTTGCTTGAAGACCGGCTACATCGGACAGTG	1488
Qy	461	Thr1yrsAspLeuLeuProTh1rAspSer1aG1nAspLeu1aSerG1uLeuVal1H1sTy1r	480
Db	1490	ACCTATATATCTGCTCCCAACGAGACAGTGCCAGACCTCGCTGCTGAATCTAGTCATAT	1548
Qy	481	GlyPheLeuH1sG1uAspAspArgMeLysLeu1a1aPheLeuG1uSerThrPheLeu	500
Db	1550	GGCTTCTCTGACGAGATGACAGGACAAAGCTAGCACCCTTCTGAGAGACACATTTCTC	1608
Qy	501	LysTy1rArgG1yThrG1n1a1a507	
Db	1610	AAGTACGAGGAGACGCAAGCG 1630	
RESULT 7			
US-10-175-523-117			
Sequence 117, Application US/10175523			
GENERAL INFORMATION: US20030096264A1			
APPLICANT: Brockman, Jeffrey			
APPLICANT: Evans, David			
APPLICANT: Hook, Derek			
APPLICANT: Klimczak, Leszek			
APPLICANT: Laeng, Pascal			
APPLICANT: Paleyeman, Michael			
APPLICANT: Rajan, Prithi			
TITLE OR INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPTS)			
FILE REFERENCE: 3335/1J795-US3			
CURRENT APPLICATION NUMBER: US/10/175, 523			
PRIORITY FILING DATE: 2002-06-18			
PRIORITY APPLICATION NUMBER: US 60/299,151			
PRIORITY FILING DATE: 2001-06-18			
PRIORITY APPLICATION NUMBER: US 60/317, 828			
PRIORITY FILING DATE: 2001-09-07			
PRIORITY APPLICATION NUMBER: US 60/325,150			
PRIORITY FILING DATE: 2001-09-25			
PRIORITY APPLICATION NUMBER: US 60/333,047			
PRIORITY FILING DATE: 2001-11-14			
PRIORITY APPLICATION NUMBER: US 60/349, 936			
PRIORITY FILING DATE: 2002-01-18			
PRIORITY APPLICATION NUMBER: US 60/361, 834			

```
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 117
; LENGTH: 3538
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-175-523-117

Alignment Scores:
Pred. No.:      1,656-276      Length:      3538
Score:          2368.00        Matches:     460
Percent Similarity: 87.17%    Conservative: 2
Best Local Similarity: 86.79%  Mismatches:  0
Query Match:    86.69%        Indels:      68
DB:             Gaps:         2

US-10-618-941-99 (1-507) x US-10-175-523-117 (1-3538)

QY      40  ArgArgGluGlnValAsnGlnGlyValMetProGlyLeuGlnSerThrPheLeuAlaMet 59
      1  CAGAGAGAGAGAGTAAACCAAGGAAACATGCCAGGAGCTTCAGAGCACTTCCTAGCCATG 60
QY      60  AaPThrgLugLugLValGluValValTPraNGluLeuNH1aPheGlyAspArgLysAla 79
      61  GACACGAGAGAGGGGGGTAGAGGTGTGTGAAACAGCTCCACTTCGAGACAGAGAGGCC 120
QY      80  PheAlaAlaNH1aGluGluVal1eGlnThrVal1PheGluGlnLeuValLeuValNH1a 99
      121  TTCGGGGGGCCAGAGAGAAAGATCCAGACCCTGTTCGAGCAGCTGTCTGGTGAGCCAC 180
QY      100 ProAlaNH1ValLeuNH1aValLeuNH1aPheNH1aPheNH1aPheNH1aPheNH1a 119
      181  CCGAACATCGTAAGTGCACAAAGTACTGCTGTGATCTCGAGCCCTGCGCGAGGCTC 240
QY      120 1lePheNH1eThrgLutYrValSerSerGlySerLeuLysGlnPheLeuLysThrLys 139
      241  ATCTTCATCAGAGATACGTGTCTCATCAGGAGCCCTCAAGCAATTCTCTAAAGACAAAG 300
QY      140  LysAlaNH1aValMetAlaNH1aNH1aNH1aNH1aNH1aNH1aNH1aNH1aNH1a 148
      301  AAGAACCACAAAGGCATGAACGCCGGGTATGGAGAGCGGCTGGGAGCAGCAGGGAGC 360
QY      148  ----- 148
      361  AGAGCGGGGTGGGAGCGCTCGGGAGCTGGAGTGTAGAGGGGTGCCCGCGGCTTCGG 420
QY      148  ----- 148
      421  ACAGGGGCTGGGCGAGAGATCGGGGGCGGCTCCGAGGCCCAAGCCGCTTCCTCGGCC 480
QY      149  ----- 149
      481  ACCCGACCGACGAGATCGTGCCTCCGCCCGCAGCCTGGAAGCGCTGGTGCACGAGATC 540
QY      158  LeuSerAlaLeuSerPheLeuNH1aLysSerProProLysLysLysLysLysLysLys 177
      541  CTGTGTGCGCTAGCTTCCTGTGACCGCTGCACGCCCAACATCATCAAGGAGAACTGACC 600
QY      178  SerAaPThrllePheNH1eGlnNH1aAsnGlyLeuLysLysLysLysLysLysLys 197
      601  AGCGACACATCTTATTCAGACCAACGGCCCTCATCAAGATGGCTCCGTGTGGACCGGA 660
QY      198 1lePheSerAsnAlaLeuArgProProThrlaLeuProAspAspLeuArgSerProLle 217
      661  ATCTCTCCCAAT-----GCACCTTCAGATGATCTCCGAAGCCCATC 702
QY      218  ArgAlaGluArgGluGluLeuArgAsnLeuNH1aPhePheProProGluTyrGlyVal 237
      703  CGCGGTGAGCGAGAGAACTTCGAACTTCCTTCCTCCCGCAGATATGAGAGGTG 762
QY      238  ALaAaPrgLutYrAlaValaPheNH1ePheSerPheGlyMetCysAlaLeuGluMetAlaVal 257
      |||||||
```

```
DB      763  GCCAGTGGAGACCGGTGTGGACATCTTCCTTGTGGAGTGTGTGGCGTGGAGATGGCTGTA 822
QY      258  LeuGluNH1eGlnThrlaNH1aPheThrArgVal1ThrgLugLual1leAlaArgAlaArg 277
      823  CTGAAATTCAGACCAATGGGGAGACCGGGGTCAAGAGGGGCATTGCTCGCGCCAGG 882
QY      278  NH1aSerLeuSerAaPProAspMetArgLusPheNH1eLeuNH1eCysValLeuAlaArgAaPPro 297
      883  CACTCGCTGAGTACCCCAACATCGGAGATTCTTCCTTGTGTGTGGCGCGGAGACCT 942
QY      298  ALaArgArgProSerAlaNH1aSerLeuLeuPheNH1aArgVal1LeuPheGluValNH1aSer 317
      943  GCCCGCGCGCCCTCTGCCCAAGCCTCTTCCACCGCGGTCTTTCAGAGTGCATCG 1002
QY      318  LeuLysLeuLeuAlaAlaNH1aCysPheNH1eGlnNH1aGlnTyrLeuMetProGluNH1aVal 337
      1003  CTGAAGCTCTCGGAGCGCCACCTGCTTATCCAGACCAAGTACCTCATGCCAGAGATGTG 1062
QY      338  ValGluGluYrThrlaValaMetAspLeuNH1aValaValaLeuAlaGluLeuProArgPro 357
      1063  GTGAGAGAGAGAACCAAGCCATGACCTGCACCGGTCTTGGCGAGCTTCGCCGGGCC 1122
QY      358  ArgArgProProLeuGlnTyrArgTyrSerGluValSerPheMetGluLeuAspLysPhe 377
      1123  CGCAGGCCCCCGCTGCAGTGGCGGTACTCGGAATCTCTTCATGAGAGCTGGACAAATTC 1182
QY      378  LeuGluAspValaArgNH1aGlyLysTyrProLeuMetAlaNH1aAlaThrArgProLeu 397
      1183  CTGAGAGATGTCAAGAAATGGAATCTACCCACTGATGAACCTTGGAGCAGCAGCCCTG 1242
QY      398  GlyLeuProArgValaLeuAlaPProProGluGluValaGluNH1aValaValaThrlaProThr 417
      1243  GGGCTGCCCGTGTGTGGCCACCCCGAGAGAGGTCCAAAGGCCCAAGACCCCGAGCG 1302
QY      418  ProGluProPheAaPSerGluThrArgLysVal1leGlnMetGlnCysAsnLeuGluArg 437
      1303  CCAAGGCCCTTTGACTGTGAGACCAAGAGTCAATCCAGATGAGTGAATCGGAGAGGA 1362
QY      438  SerGluAspLysAlaArgTyrPheNH1eLeuThrlaLeuLeuVal1leGluAspArgLysNH1a 457
      1363  AGCAGAGACAAAGGCCCTGTCATCTGCTTGTGTGGTGGAGAGCGGCTGTGCAC 1422
QY      458  ArgGlnLeuThrlaPheLeuLeuProThrlaPheSerAlaGlnaPheLeuAlaSerGluLeu 477
      1423  CGGAGCTGACCTTACAGCTGTCTCCAAAGGACAGCGCCAGACCTCGGCTCGAGAGTC 1482
QY      478  ValNH1aTyrGlyPheLeuNH1aGluAspAspArgMetLysLeuAlaAlaPheLeuGluSer 497
      1483  GTGCATATGGCTTCTCTCAAGAGACGACCGGATGAAGCTGGCGGCTCTGGAAGAGC 1542
QY      498  ThrPheLeuLysTyrArgGlyThrlaAla 507
      1543  ACCTTCCTCAAGTACCGTGGGACCCAGGCC 1572
DB

RESULT 8
US-11-099-266-117
; Sequence 117, Application US/11099266
; Publication No. US20050181433A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Paley, Pascal
; APPLICANT: Rajan, Priti
; TITLE OR INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 03235/100795-US4
; CURRENT APPLICATION NUMBER: US/11/099,266
; PRIOR FILING DATE: 2005-04-04
; CURRENT APPLICATION NUMBER: US 10/175,523
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
```

/ PRIOR FILING DATE: 2001-06-18
/ PRIOR APPLICATION NUMBER: US 60/317,828
/ PRIOR FILING DATE: 2001-09-07
/ PRIOR APPLICATION NUMBER: US 60/325,150
/ PRIOR FILING DATE: 2001-09-25
/ PRIOR APPLICATION NUMBER: US 60/333,047
/ PRIOR FILING DATE: 2001-11-14
/ PRIOR APPLICATION NUMBER: US 60/349,936
/ PRIOR FILING DATE: 2002-01-18
/ PRIOR APPLICATION NUMBER: US 60/361,834
/ PRIOR FILING DATE: 2002-03-04
/ NUMBER OF SEQ ID NOS: 197
/ SOFTWARE: Patentin version 3.1
/ SEQ ID NO 117
/ LENGTH: 3538
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-11-099-266-117

Alignment Scores:
Pred. No.: 1,65e-276 Length: 3538
Score: 2368.00 Matches: 460
Percent Similarity: 87.17% Conservative: 2
Best Local Similarity: 86.79% Mismatches: 0
Query Match: 88.69% Indels: 68
DB: 10 Gaps: 2

US-10-618-941-99 (1-507) x US-11-099-266-117 (1-3538)

QY 40 ArgatguglgnlvalaanglnglyasnmecProglYleuGlnserThrpheleuAlamet 59
Db 1 CAGAGGAGAGAGTAAACCAAGGAAACATGCCAGGCTTCAGAGCACTTCTTAAGCAAG 60
QY 60 AaprrtGluGluGlyValaGluValaITTPraanglYleuHlphelGlyAaprrglYala 79
Db 61 GACAGGAGAGAGGAGGTAGAGGTGTGTGAAACAGAGCTCAGTTCGAGAGCAAGAGGCT 120
QY 80 PheAlaAlaHlGluGluYValleGlnThrpValPheGlnGluValleuValaAaprrHs 99
Db 121 TTCGGGAGCCAGAGAGAGAAAGATCCAGAGCTGTGTGACAGCTGTGTGTGAGACAC 180
QY 100 ProaenlleValYleuHlshYsYrTTPleuAaprrThSerGluAlaCyaaAlaArgVal 119
Db 181 CCGAACATGTGAAGTTCAGAAAGTACGTGGTGAATCCTTCAGAGCTTCGCGAGGCTC 240
QY 120 HlphelIethrGluYrValserSerGlySerleuYglnPhelYleuYshYrThrlYs 139
Db 241 ATCTTCATCAGAGAGTACGTGTATCAGAGCAGCTCAAGCAATTCCTCAAAAAGACAAAG 300
QY 140 LyaaenHlshYAlaHlshYAlaHlshYAlaHlshYAlaHlshYAlaHlshYAlaHlshY 148
Db 301 AAGAACCAACAGGCAATGAACGCCCGGTATGTGGGAGCGGCTGTGGGAGCCAGCGGAG 360
QY 148 ----- 148
Db 361 AGGAGGGGTTGGGCGAGCTCGGGAGCTGGAGTGTGAAGGGGTGCCCGCGGCTCGG 420
QY 148 ----- 148
Db 421 ACAGGGGCTGGCGAGATCGGGGCGGCTCGGAGGCCAGCGGCTTCCTGGGCCC 480
QY 149 ----- 148
Db 481 ACCCGACCGAGAGT 540
QY 158 LeuSerAlaLeuSerPheleuHlshYAlaCySerProProIleHlshYAlaHlshYAlaHlshY 177
Db 541 CTGTCTCGGCTCAGCTTCTGTGACGCTGTGACGCTGTGACGCTGTGACGCTGTGACGCTG 600
QY 178 SerAaprrThrlPheHlshYAlaHlshYAlaHlshYAlaHlshYAlaHlshYAlaHlshY 197
Db 601 AGCGACACCATCTTCATTCAGACAAAGGCTCATCAAGATCGGCTCGGTGTGAGACCGA 660

QY 198 HlpheserAaAlaLeuArgProProThrlAlaLeuProAaAaprrleuArgSerProIle 217
Db 661 ATCTTCTCAAT-----GCACTTCAGATGATCTCGAAGCCCCATC 702
QY 218 ArgAlaGluArgGluGluYleuAArgAenHlshYpHePheProProGluYrGluYVal 237
Db 703 CGCGGTGAGCAGAGAGAACTTCGAAACCTGCACTTCTTCCCGCAGAGATGAGAGCGTG 762
QY 238 AlaAaprrGlyThAlaValAaprrlePheSerPheGlyMetCyAlaLeuGluMetAlaVal 257
Db 763 GCCGATGGGACCGCTGTGACATCTTCTTGGGATGTGTGCGGTGAGATGGCTGTGA 822
QY 258 LeuGluIleGlnThrpAanglYasprHrArgValITPnglGluAlaIleAlaArgAlaArg 277
Db 823 CTGAATATCAGACCAATGGGAGACCGGGGTACAGAGAGGCAATGCTCGCGCAGG 882
QY 278 HlsserleuSerAaprrProAaHlshYArgGluPheleleuCyCySerleuAlaArgAaprr 297
Db 883 CACTCGCTGAGTGAACCCCAACATGGGGAGTTCATCTTGTGCTGGCCCGGAGCCTT 942
QY 298 AlaAargProSerAlaHlsserleuPheHlshYValleuPheGluValHlsser 317
Db 943 GCCCGCCGCTCTGTGCCCACAGCTCTTCCACCGGCTCTTTCAGAGGTGACACTCG 1002
QY 318 LeuYshYleuAlaAlaHlshYpHeHlshYAlaHlshYAlaHlshYAlaHlshYAlaHlshY 337
Db 1003 CTGAAGCTCTGGAGCCCACTGTTCATTCAGACACAGTACTCTGCTGAGAAATGTG 1062
QY 338 ValGluGluYrThrYshYAlaHlshYpHeHlshYAlaValleuAlaGluYleuProArgPro 357
Db 1063 GTGAGAGAGAGAACCAAGGCCATGAGCTGTGACGCGGCTTGTGGAGGTTCCCGGCCC 1122
QY 358 ArgatgProProleuGlnTTPargYrSerGluValSerPheMetGluYleuAaprrYsh 377
Db 1123 CGCAGGCCCCGCTCAGTGGGCTGTATCTGGAAGCTCTTCATGAGAGTGAACAAATTC 1182
QY 378 LeuGluAaprrValArgAanglYrIleYrProleuMetAaPheAlaAlaThrlArgProleu 397
Db 1183 CTGGAGAGATGTCAAGAAATGAATCTAACCACTGAAGAACTTTGACAGCACTGACACCTCG 1242
QY 398 GYleuProArgValleuAlaProProProGluGluValaGlnYshYAlaYshYrProThr 417
Db 1243 GGGCTGCCCGGTCTGTGCCCCACCCCGAGAGGTGTCAAAAGGCCAAGCCCCGAGG 1302
QY 418 ProGluProPheAaprrSerGluThrlArgYshYAlaIleGlnMetGlnCyAaHlshYAla 437
Db 1303 CAGAGGCCCTTGAATCTGAGACCAAGAAAGTCAATCCAGATGCAATGCACTGAGAGA 1362
QY 438 SerGluAaprrYshYAlaArgTThrlshYleuThrlleuYleuValleuGluAaprrYshYsh 457
Db 1363 AGCGAGACCAAGGCGCGGTGATCTCACTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1422
QY 458 ArgGlnleuThrlYrAaprrleuYleuProThrpAaSerAlaGlnAaprrleuAlaSerGluYleu 477
Db 1423 CGGAGCTGACCTTACAGCTGTCTCCAAAGAGACGCCCAAGAGCTTCCTCGAGGCTC 1482
QY 478 ValHlshYrGlyPheleuHlshYAlaAaprrArgMetYshYleuAlaAlaPheleuGluSer 497
Db 1483 GTGACATATGGCTTCTCAAGAGGACCAAGGATGAAGTGTGGCGGCTTCCTGAGAGGC 1542
QY 498 ThrlPheleuYshYrArgGlyThrlGlnAla 507
Db 1543 ACCTTCTCAAGTACCGTGGAGCCAGCC 1572

RESULT 9
US-09-862-027-9
; Sequence 9, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027


```

Db      275 GGGGCGCAGAGAGATCCAGACCGATTGTCAGACAGCTGCTGTCGAGCACCAG 334
Qy      101 Aenllvalylleuuhlsyetyrttpleuapthserglualacyalaargvalille 120
Db      335 AACATCGTAAAGTTGACAGAGTACTGGCTGGAACTCTGAGAGCCCTGCGAGGGTCAATC 394
Qy      121 Pheilethglutylvalsersergllyserleuylleuylpheuueylserthrllyslvs 140
Db      395 TTCAATCAGAGTACGTCTCATTCAGCAGCCTTCAAGCAATTCCTCAAAAAGACCAAGAG 454
Qy      141 AannlsylvalamerannalaargalatrplysartrpCyethrlnlleuSerala 160
Db      455 AACCAAGGCGCATGAAACCGCGGCTGGAAGCGCTGGAGCAGCAGATTCGTGTGG 514
Qy      161 leuSerPheleuHlsalacySerProProllelHlsiglYlanleuThrsarPthr 180
Db      515 CTCAGCTTCTGCACCCCTGCAGCCCCCAATCATCCAGGGAACTTGACCAAGCCAGCAC 574
Qy      181 llePheileglHlsanlglYleuileyllelglYserValThrlsarPthlePser 200
Db      575 ATCTTCATTCAGCACAAGCCTTCATCAAGATCGCTCCGTGGCAGCAATCTTCTCC 634
Qy      201 AannalaleuargProProthralaleuProAapPleuargSerProilearagalgu 220
Db      635 AAT-----GCACCTTCAGATGATCTCCGAAGCCCCATCCGCGCTGAG 676
Qy      221 ArgglugluYleuarglanleuHlsPhePheProProlguYrlygluValalaargly 240
Db      677 CAGAGAGAACTTCGGAACCTGCACCTTCCCCCAGAGTATGAGAGGTGGCCGATGG 736
Qy      241 ThrAlaValAspIlePheSerPheglYmetCybalaleuGlumetAlaValleuGuile 260
Db      737 ACCGCTGTCGACATCTTCTTTGGATGTGCGCTGAGATGGCTGTACTGTGAATC 796
Qy      261 GlnThraenlYAsPtharYValThrglugluAlaAlaArgAlaargHlsSerleu 280
Db      797 CAGACCATGGGAGACACCGGCTCAGAGAGGCAATGCTCGCCCGGCACTGCGT 856
Qy      281 SerAapProAmetarggluPheileuYsCyaleuAlaargAapProAlaargarg 300
Db      857 AGTGACCCCAACATCGGAGATTCACTTGTGCTGCGGAGCCCTCGCCGCGC 916
Qy      301 ProSarAlaHlsSerleuPheHls-ArgValleuPhegluValHlsSerleuYale 320
Db      917 CCTCTGTCACAGCTCTCTTCCACNCGCGTCTTNGAGGTGCACCTCGTAAAGCT 976
Qy      320 uleu 321
Db      977 CCG 980

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/840,787
FILING DATE: 23-Apr-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/518,865
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0356 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 69:
SEQUENCE CHARACTERISTICS:
LENGTH: 2029 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAITVT03
CLONE: 864683
SEQUENCE DESCRIPTION: SEQ ID NO: 69 :
US-09-840-787-69
US-10-618-941-99 (1-507) x US-09-840-787-69 (1-2029)
Alignment Scores:
Pred. No.: 2,266-175
Score: 1534.50
Percent Similarity: 73.54%
Best Local Similarity: 61.62%
Query Match: 57.47%
DB: 3 Gaps: 6
Qy      17 GIUARGUAspGluSerGluAspGluSerAplleuGlulSerProcyeglyarg 36
Db      240 GAGGAAGAAAGAAAGAGATGATGATCTGAGATTTGGAAAGATCGCCCTGTGGCC 299
Qy      37 TTPGlnlysarYarggluGluValaenlglYlanmetProglYleuGlnSerthrphe 56
Db      300 TGGCAAGAGGCGGAAAGAGTGAATCAACGAAATGTAACAGATTTACAGTGCATAC 359
Qy      57 leuAlametAapThrglugluYValgluValaITTPAenGluLeuHlsPheglYAsp 76
Db      360 CTGGCCATGATATACAGAGAGAGTGTAGAGTTGTGGAAATGAGGTACAGTTCTGAA 419
Qy      77 ArglyValaPheAlaHlsiglYleuYslleGlnThrValPhegluGlnleuValleu 96
Db      420 CGCAAGAACTACMACTGCAGAGGAGAAAGTTCTGCTGTGTGTGATATCTGATTCAA 479
Qy      97 ValaPheIspProAenlleVallyslleuHlsyetyrttpleuapthserglualacy 116
Db      480 TTGAGACATCTTAACATTTGTTTCAAAATTTGGGCTGACATTTAAAGAAACAG 539
Qy      117 AlaargValallePheilethrgluYrlyValserserglYserleuylleuPheleuY 136
Db      540 GCCAGGTCATTTATACAGAAATACATGTCATCTGGAGAGTGAAGCAATTTCTGAAG 599
Qy      137 lystrlyslvalaenHlsylalametannalaargalatrplysartrpCyethrln 156
Db      600 AAGACCAAAAGAAACCAAGACGATGATGAATGAAGGAGATGAAACCTTGGTGCAACA 659
Qy      157 lleuSerAlaIleuSerPheleuHlsalacySerProProllelHlsiglYlanleu 176
Db      660 ATCTCTGCTGCTTAAGTACTGCACTCTGTCAGTCTGTCAGTCTGTCAGTCTGTCAGT 719

```

```

QY 177 ThrsrAaPThrIlePheIleGlnHisAengIleuIIeIyIleGlySerValTrrHis 196
DB 720 ACCTGTGACACCATTTCTATCCAGCAACAGACATCATCAATGATGCTGTG----- 773
QY 197 ArgIlePheSerAaAlaLeuArgProProThrAlaLeuProAaPLeuArgSerPro 216
DB 774 -----GCTCCCTGACACATCAACATCAT 797
QY 217 IleArgAlaGluArgGluGluLeuArgAenLeuHlePhePheProProGluTyrGlu 236
DB 798 GTGAAGACTTGTGAGAAAGACAGAAATCTACACTTCTTGGACACAGATATGAGAGA 857
QY 237 ValAlaAaP---GlyThrAlaValAerIlePheSerPheGlyMetCysAlaLeuGlnMet 255
DB 858 GTCACTAATGTGACAAACAGACATGACATCTCTCTTGGCATGTGTGACATGAGATG 917
QY 256 AlaValLeuGluIleGlnThrAaengIyAerThrArg---ValThrGluGluAlaIleAla 274
DB 918 GCAGTGTGTGAGATTCAGGGCAATGAGAGTCTCATATGTCCACAGAGCCATCAAGC 977
QY 275 ArgAlaArgHisSerLeuSerAaPProAaMetArgGluPheIleLeuCysCysLeuAla 294
DB 978 AGTGCATCCAGCTTCTAGAAAGACCCATTACAGAGGAGTTCATTCAAAAGTGCCTGCAG 1037
QY 295 ArgAaPProAlaArgArgProSerAlaHisSerLeuPheHisArgValLeuPheGlu 314
DB 1038 TCTGAGCTGTGTGAGACAAACAGCAGAGAACTTGTGTCCACCCAGCATTTGTTGAA 1097
QY 315 ValHisSerLeuLeuLeuAlaAlaHisCysPheIleGlnIleGlnTyrLeuMetPro 334
DB 1098 GTGCCCTGTGCTCAAACTCTTGGGCCCATCTGCATTTGGGACAAACACATGATCCCA 1157
QY 335 GluAaenValValGluGluLeuThrIyArgAlaMetAerPheHisAlaValLeuAlaGluLeu 354
DB 1158 GAGAAACGCTGTAGAGAGATCAACAAAAACATGATACTAGTACGCGTGTGCTGAATTC 1217
QY 355 ProArg---ProArgArgProProLeuGlnTyrArgTyrSerGluValSerPheMetGlu 373
DB 1218 CTTGCAGAGCCAGAAAGAACCCAGTTCCAGACTTGTACTCTCACTCAACAGCTTGGA 1277
QY 374 LeuAerPhePheLeuGluAerValArgAaengIyIleTyrProLeuMetAaPheAlaAla 393
DB 1278 TTAGTAAATTCCTTGAAGATGTCAAGAAATGGATCTATCTCTCAAGCCCTT----- 1331
QY 394 ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal----- 410
DB 1332 -----GGGCTGCTCGGCCCCAGCAGCAGCAGAGAGAGAGTGCATCATCCT 1379
QY 411 -----GlnIyAlaIleThrProThrProGluProPheAaPserGluThrArgIyS 427
DB 1380 GTCTGCCCCCTGTGTCAAGCTCCGACCTGAAACAGCTGAGTGAAGCTGCAAG 1439
QY 428 ValIleGlnMetGlnCysAaenLeuGluArgSerGluAerPheAlaArgTrrHisLeuThr 447
DB 1440 GTGGTGTGATTCGATGCAACATTAAGTCGGTGGAGAGAGAGTCAAAACACACCTGACA 1499
QY 448 LeuLeuLeuValLeuGluAerArgLeuHisArgGlnLeuThrTyrAaPLeuLeuProThr 467
DB 1500 CTTTGTGTAAGTGTGAGAGACAACTGAACCGGACCTGAGCTGAGCTGAGCTGAGCAAT 1559
QY 468 AaPserAlaGlnAaPLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAaPLeu 487
DB 1560 GAGAAATATCCCGAGTGTGGCTGAGCTGTGAGCTGTGGCTTATTATGAGAGCTGAC 1619
QY 488 ArgMetIyPheValAlaPheLeuGluSerThrPheLeuIyTyr 502
DB 1620 CAGAGCCGGTTCCTTCTGTCTAGAAAGACCTTGAACAAAGTTC 1664

```

```

; GENERAL INFORMATION:
; APPLICANT: HONDA, Goichi
; APPLICANT: MATSUDA, Akio
; APPLICANT: MURAMATSU, Shuji
; APPLICANT: ISHIZAWA, Kenya
; TITLE OF INVENTION: STAT6 Activating Gene
; FILE REFERENCE: 1254-0207P
; CURRENT APPLICATION NUMBER: US/10/153,668
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 60/293,172
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/316,031
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/328,403
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: JP 2001-157043
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: JP 2001-260681
; PRIOR FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: JP 2001-313175
; PRIOR FILING DATE: 2001-10-10
; NUMBER OF SEQ ID NOS: 488
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 321
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (113)..(1717)
US-10-153-668-321

Alignment Scores:
Pred. No.: 2,5e-175 Length: 2163
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: 5 Gaps: 6

US-10-618-941-99 (1-507) x US-10-153-668-321 (1-2163)
QY 17 GluArgGluAaPLeuGluSerGluAaPLeuGluSerAerIleLeuGluGluSerProCysGlyArg 36
DB 242 GAGAAAGAAAGAAAGAGATGAGTGTGAGATTTTGGAAAGTGCCTGTGTGGCGC 301
QY 37 TyrGlnIyAaArgAaGluGluGlnValAaengIyAaenMetProGlyLeuGlnSerThrPhe 56
DB 302 TGGCAAGAAAGCCAGAAAGAGTGAATCAAGGAATGTAACAGATATTGACAGTGCATAC 361
QY 57 LeuAlaMetAaPThrGluGluGlyValGluValAlaTrrAaengIyLeuHisAerGlyAaP 76
DB 362 CTTGCCATGATTAACAGAGAAAGTGTGAGATTTGTGAATGAGTACAGTCTCTGAA 421
QY 77 ArgIyAaPheAlaAlaHisGluGluLeuIleGlnTrrValPheGlnLeuValLeu 96
DB 422 CGCAAGAACTACAGAGCTGCAGAGAAAGGATGTGTGTGTTGAATTAATCTGATTCAA 481
QY 97 ValAaPHisProAaenIleValIyLeuHisIyTrrTrrLeuAaPThrSerGluAlaCys 116
DB 482 TTGAGAGCATTTAATCATTTGTTAAAGTTTCAACAAATTTGGCTGACATTAAGAAACAAG 541
QY 117 AlaArgValIlePheIleThrGluTyrValSerSerGlySerLeuIyGlnPheLeuIyS 136
DB 542 GCCAGGCTCATTTTATTCACAGAAATACATGTCATCTGGAGTGCAGCAATTTCTGAAG 601
QY 137 LyeThrIyAaPLeuAaPLeuHisIyValMetAaPAlaArgAlaTrrIyAaPLeuTrrCysThrGln 156
DB 602 AAGACCAAAAGAAACCAAGACAGATGAATGAAGAGATGAAGCGTTGGTGCACAA 661
QY 157 IleuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAaenLeu 176
DB 662 ATCTCTCTGCTTAAGTACTGACTCTCTGTGAACCCCAATCATTCAGTGAACCTG 721

```

177 ThrSerAspThrIlePheIleGlnIleAsnGlyLeuIleValIleGlySerValTrpHis 196
172 ACCGTGAGACCATCTTCAATCCAGCAACAGCACTATCAAGATTGGCTGTGG----- 775
197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
776 -----GCTCCCTGACACTTCAACATCAT 799
217 IleArgAlaGluArgGluGluLeuArgAsnLeuIlePhePheProProGluTrpGlyGlu 236
800 GTGAAGACTTGTCGAGAAAGAGCAAGAAATCTACCTCTTGGCCACAGAGTATGAGAGA 859
237 ValAlaAsp---GlyThrAlaValAspIlePhePheSerPheGlyMetCysAlaLeuGlnMet 255
860 GTCACTATGTGACAAACAGAGTGCACATCTACTCTTGGCAGTGTGTGCATGGAGATG 919
256 AlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla 274
920 GCAGTGTGAGATTGAGGGCAATGAGAGTCTCTCATATGTGCCACAGAAAGCCATCAGC 979
275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysCysLeuAla 294
980 AGTGCATCAGCTTCTTAGAAGCCCATTAACAGAGGAGTTCAATCAAAAGTCCCTGAG 1039
295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
1040 TCTGAGCTGTCTGCAGACCAACAGCAAGAACTTCTGTCCACCCAGCATTTGTTGAA 1099
315 ValHisSerLeuLeuLeuAlaAlaHisCysPheIleGlnHisGlnTrpLeuMetPro 334
1100 GTGGCTGCTCAAACTCTTGGGCGCCACTGTGATTTGGAGACCAACACATGATCCCA 1159
335 GluAsnValValGluGluLeuThrIleValAspLeuHisAlaValLeuAlaGluLeu 354
1160 GAGAACTGTCTAGAGAGATCAACAAACATGAGATACATGAGTCCGTATGCGCTGAATC 1219
355 ProArg---ProArgArgProProLeuGlnTrpArgTrpSerGluValSerPheMetGlu 373
1220 CTTGCAGAGCCAGAAAGAAACAGATTCAGACTTGTACTCTCAGCCACAGCTCTGAGA 1279
374 LeuAspLeuPheLeuGluLeuAspValArgAsnGlyIleTrpProLeuMetAsnPheAla 393
1280 TTAGATTAATTTCTTGAAGATGTCAAGAAATGGGATTAATCTCTGACACCTTT----- 1333
394 ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal----- 410
1334 -----GGGCTGCTCGGCGCCAGAGCCAGACAGAGAGAGGATGATCACCCT 1381
411 -----GlnIleValAlaValSerThrProThrProGluProPheAspSerGluThrArgGly 427
1392 GTCTGCCCCCTCTGTCAAGACTCGAGACCTGAACCTGAGCTGAGTGAAGCTCGCAAG 1441
428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLeuAlaArgTrpHisLeuThr 447
1442 GTGGTGTGATGCAAGTGAACATTAGTGTGAGAGAGAGAGTCAACACCACTGAGAA 1501
448 LeuLeuLeuValLeuGluLeuAspArgLeuHisArgGlnLeuThrTrpAspLeuProThr 467
1502 CTTCTGCTGAAGTTGAGAGCAAACTGAACCGGACCTGAGCTGTGACTGATGGCAAT 1561
468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisTrpGlyPheLeuHisGluAspAsp 487
1562 GAGAAATATCCCGAGTGTGGCTGTGAGTGTGAGCTGGGCTTCAATGTAGAGCTGAG 1621
488 ArgMetCysLeuAlaAlaPheLeuGluSerThrPheLeuValTrp 502
1622 CAGAGCCGGTTGACTTCTCTGCTAAGAGAGACTTGAACAGATTTC 1666

RESULT 13
US-10-887-553A--882
; Sequence 882, Application US/10887553A
; Publication No. US20050085436A1

GENERAL INFORMATION:
APPLICANT: Garza, Dan
TITLE OF INVENTION: Method to treat conditions associated
FILE REFERENCE: 4-33262
CURRENT APPLICATION NUMBER: US/10/887,553A
PRIOR APPLICATION NUMBER: 60/485,883
PRIOR FILING DATE: 2003-08-07
NUMBER OF SEQ ID NOS: 1208
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 882
LENGTH: 2163
TYPE: DNA
ORGANISM: human
US-10-887-553A-882

Alignment Scores:
Pred. No.: 2,5e-175 Length: 2163
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: Gaps: 6

US-10-618-941-99 (1-507) x US-10-887-553A-882 (1-2163)

17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArg 36
242 GAGGAAGAAAGAAAGAAAGTAAGTGAATGATTTGGAAAGTCCCTGTGGCGCC 301
37 TrpGlnLeuArgArgGluGluValAsnGlnIleAsnMetProGlyLeuGlnSerThrPhe 56
302 TGGCAGAAAGCGCGAAGAGGTGAATCAACGAAATGTACAGGTATGACAGTGCATAC 361
57 LeuAlaMetAspThrGluGluGluValGluValAlaTrpAsnGluLeuHisPheGlyAsp 76
362 CTGGCCATGATTAACAGAGAGAGGTGAGGTGTGTGGAAGAGGTACAGTCTTGAA 421
77 ArgIleValAlaAlaHisGluGluValIleGlnThrValPheGluGlnLeuValLeu 96
422 CCGAAGACTACAGCTGAGAGAGAAAGTTCGCTGTGTTCATATCTGATTCAA 481
97 ValAspHisProAsnIleValIleValLeuHisValTrpLeuAspThrSerGluAlaCys 116
482 TTGAGACATCTTAACATTTGTAAGTTTCAAAATATTTGGCTGACATTAAGCAAG 541
117 AlaArgValIlePheIleThrGluTrpValSerSerGlySerLeuValGlnPheLeuVal 136
542 GCCAGGGTCATTTTATCAACAGAAATACATGATCTGGAGAGTCTGAAGCAATTTCTGAG 601
137 LysThrIleValAsnHisIleValAlaMetAsnAlaAspAlaTrpLysArgTrpCysThrGln 156
602 AAGACCAAAAGAACCAACAGACATGAATGAAGGACATGAAGGCTGTGGCACACAA 661
157 IleuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
662 ATCCCTCTGCCCTTAAGTACTGCACTCTGTGACCCCTCATCATCATGGAACCTG 721
177 ThrSerAspThrIlePheIleGlnHisAsnGlyLeuIleValIleGlySerValTrpHis 196
722 ACCCTGTACACCATCTTCAATCCAGCAACAGACATCATCAATGATGGCTGTG----- 775
197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
776 -----GCTCCCTGACACTTCAACATCAT 799
217 IleArgAlaGluArgGluGluLeuArgAsnLeuIlePhePheProProGluTrpGlyGlu 236
800 GTGAAGACTTGTCGAGAAAGAGCAAGAAATCTACCTTCTTGGCCACAGAGTATGAGAGA 859
237 ValAlaAsp---GlyThrAlaValAspIlePhePheSerPheGlyMetCysAlaLeuGlnMet 255

Db 1100 GTGCCCTCCGCTCAAACTCCTTGGCGCCACTGCATTGTGGACACCAACATGATGCCA 1159
Qy 335 GUAAATValAGluGluValThrValMetAspLeuHisAlaValLeuAlaGluLeu 354
Db 1160 GAGAAAGCTCTAAGAGATGATCAACCAAAACATGATATAGTCCCTGACTGGCTGAAATC 1219
Qy 355 ProArg---ProArgArgProProLeuGlnTTPArgTyrSerGluValSerPheMetGlu 373
Db 1220 CTTGCAGAGACCAAGAAAGAACAGTTCACACTTTGACTCTCAGTCACCAAGCTTGGAA 1279
Qy 374 LeuAspLysPheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 1280 TTAGATTAATTCCTTGAAGATGTCAGAAATGGATATATCTCTGCACAGCTTT----- 1333
Qy 394 ThrArgProLeuGluLysProArgValLeuAlaProProGluGluVal----- 410
Db 1334 -----GGGCTGCTCCGCGCCCGCCAGACCCACAGCAGAGAGGATGATCATCAGCT 1381
Qy 411 -----GlnValAlaValThrProThrProGluProPheAspSerGluThrArgLys 427
Db 1382 GTCCGCCCCCTCTCTCAAGACTCCGACACTGAAACAGCTGAGTGAAGACTGCGAAAG 1441
Qy 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLysAlaArgTyrHisLeuThr 447
Db 1442 GTGGTGCTGATGAGTGAACATATGAGTGGTGGAGAGAGAGATCAACACCACTGACA 1501
Qy 448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr 467
Db 1502 CTTCTGCTGAATGGAGAGCAAACTGAACCGGCACCTGAGCTGTGACTGATGCGCAAT 1561
Qy 468 AspSerAlaGlnAspLeuLysSerGluLeuValHisTyrGlyPheLeuHisGluAspArg 487
Db 1562 GAGAAATATCCCGAAGTTGGCGGCTGAGCTGTGACAGCTGGGCTTCAATGATGAGCTGAC 1621
Qy 488 ArgMetLysLeuAlaLysPheLeuGlnSerThrPheLeuLysTyr 502
Db 1622 CAGAGCCGGTTGACTTCTGCTGAGAAAGACCTTGAACAAGTTTC 1666
RESULT 15
US-09-764-875-151
; Sequence 151, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ202
; CURRENT APPLICATION NUMBER: US/09/764,875
; PRIORITY FILING DATE: 2001-01-17
; Prior application data removed - consult PAMM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 151
; LENGTH: 2225
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-875-151
Alignment Scores:
Pred. No.: 2,61e-175 Length: 2225
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatches: 102
Query Match: 57.47% Indels: 29
DB: 3 Gaps: 6
US-10-618-941-99 (1-507) x US-09-764-875-151 (1-2225)
Qy 17 GUAAAGLAspGluSerGluAspGluSerAspIleLeuGlnGluSerProCysGlyArg 36
Db 270 GAGGAAGAAGAAAGTGAAGATGATGTTGAGATTTTGGAAAGATGCGCCCTGTGGGCGC 329
Qy 37 TrpGlnLysArgArgGluGlnValAsnGlnLysAsnMetProGlyLeuGlnSerThrPhe 56

Db 330 TGGCAGAAAGCGAGAAAGAGATGATCAACGGAATGTACAGATTTTACAGTGCATAC 389
Qy 57 LeuAlaMetAspThrGluGluGlyValGluValValTrpAsnGluLeuHisAspGlyAsp 76
Db 390 CTGGCCATGATATCAGAGAGAGGTGTAGAGTTGTGTGGAAATGAGATGATGATCTCTGAA 449
Qy 77 ArgLysAlaPheAlaAlaHisGluGluLysIleGlnThrValPheGluGlnLeuValLeu 96
Db 450 CGCAAAATACAAAGCTGACAGAGAAAGTTGCTGTGTGTATATATCTGATTCAA 509
Qy 97 ValAspHisProAsnIleValLysLeuHisLysTyrTyrTrpLeuAspThrSerGluAlaCys 116
Db 510 TTGAGACATCTTAACATTTGTAATGTTTCAAAATATTGGCTGACATTAAAGAAACAG 569
Qy 117 AlaArgValIlePheIleThrGluTyrValSerSerGlySerLeuLysGlnPheLeuLys 136
Db 570 GCGAGGGTCAATTTTATCAGAAATATCATGTATCTGGAGTGTGAAGCAATTTCTGAAG 629
Qy 137 LysThrLysLysAsnHisLysValMetAsnAlaArgAlaIleTrpLysArgTyrCysThrGln 156
Db 630 AAGACCAAAAGAAACCAAGAGATGAAGAAAGCATGMAAGCGTTGGTGCACAA 689
Qy 157 IleLeuSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
Db 690 ATCTCTTCTGCTTAAGCTTACCTGACCTCTGTGACCCCTCATTCATGAGAACTTG 749
Qy 177 ThrSerAspThrIlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValTrpHis 196
Db 750 ACCTGTGACACCATTTTCATCCAGCAACAGGATCATGAAATGGCTGTG----- 803
Qy 197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
Db 804 -----GCTCTGACACTGTCAACAAATCAT 827
Qy 217 IleArgAlaGluValArgGluLeuArgAsnLeuHisAspPheProProGluTyrGlyGlu 236
Db 828 GTMAAGACTTGTGAGAAAGAGAGAAATGTACACTTTTGTGACACCAAGATGTGGAGAA 887
Qy 237 ValAlaAsp---GlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGlnMet 255
Db 888 GTCACTAATGTGACAAACAGCACTGACATCTATCTTGGAGATGTGTCACTGGAGAG 947
Qy 256 AlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla 274
Db 948 GCAGTGCTGAGATTCAGGGCAATGAGAGCTCATATGTCCACAGAGAACCATACAGC 1007
Qy 275 ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysCysLeuAla 294
Db 1008 AGTGCATTCACGACTTCTAGAAAGCCATTACAGAGGAGATTCAATCAAAAGTCCGAG 1067
Qy 295 ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu 314
Db 1068 TTCTAGACCTGTGCTCGAGACCAACAGCCAGAGAACTTGTGTCCACCCAGCATTTTGA 1127
Qy 315 ValHisSerLeuLysLeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetPro 334
Db 1128 GTGCCCTGCTCAAACTCCTTGGCGCCACTGCATTTGGAGACACCAACATGATCCCA 1187
Qy 335 GUAAATValAGluGluValThrValMetAspLeuHisAlaValLeuAlaGluLeu 354
Db 1188 GAGAAAGCTCTAAGAGATGATCAACCAAAACATGATATAGTCCGCTGAGTGGTGAATC 1247
Qy 355 ProArg---ProArgArgProProLeuGlnTTPArgTyrSerGluValSerPheMetGlu 373
Db 1248 CTTGCAGAGACCAAGAAAGAACCACTTCAAGATTTGACTCTCAGTACCAAGCTTGGAA 1307
Qy 374 LeuAspLysPheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 1308 TTAGATTAATTCCTTGAAGATGTCAGAAATGGATATATCTCTGCACAGCTTT----- 1361
Qy 394 ThrArgProLeuGluLysProArgValLeuAlaProProGluGluVal----- 410

```
Db      1362 -----GGGCTGCTCGGCCAGCCAGCCAGCAGAGAGGTGACATCACCT 1409
Qy      411 -----GlnlysaAlaIythrProthrProgluProphaspSergIuthrArglys 427
Db      1410 GTCGTGCCCCCTCTGTCAAGACTCCGACACTGAACCACTGAGGTGAGACTCGCAAG 1469
Qy      428 ValIleGIInMetGlnCysAsnleuGluAArgserGluAspIyAlaArgTrpHisleuthr 447
Db      1470 GTGGTGTGATGCACTGCACTTGTAGTCGGTGGAGGAGGAGTCAAAACACCCTGACA 1529
Qy      448 LeuleuleuValleuGluAspArgleuHisArgGlnleuthrTyrAspIleuProthr 467
Db      1530 CTTCTGTGAAGTTGAGAGAACAACTGAACCGGCACCTGAGCTGTGATGCCAAT 1589
Qy      468 AspSerAlaGlnAspIleuAlaSerGluIleuValHisTyrGlyPheleuHisGluAspAsp 487
Db      1590 GAGAAATATCCCGAGTTGCGCGCTGAGCTGGTGCAGCTGGGCTTCATTAGTAGGCTGAC 1649
Qy      488 ArgMetIySleuAlaAlaPheleuGluSerThrPheleuIySTyr 502
Db      1650 CAGAGCCGGTTGACTTCTCTGTAGAGAGACCTTGAAACAAGTTC 1694
```

Search completed: January 14, 2006, 03:03:27

Job time : 996 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 13, 2006, 20:06:05 ; Search time 742 Seconds
(without alignments)
4553.907 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAPEPAPRARERERERED.....RMKLAFLSEFLKXGTGTA 507

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_model -DEV=xlp
-Q=/cgn2.1/USPTO_epool_p/US10619941/runat_12012006_070203_14258/app_query.fasta_1.647
-DB=N Geneseg -QFMT=fastap -SUFFIX=ing -MIMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human0.cdi -LIST=45
-DOCALLING=200 -THR SCORE=pcpt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pcpt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10618941@cgn_1_1_1096@runat_12012006_070203_14258 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2670	100.0	3765	12	ADJ96576 Human N1m
2	2612	97.8	4048	12	ADL13324 Human N1m
3	2584	96.8	3837	13	ADK52130 Human N1m
4	2584	96.8	3838	13	ADK52108 Human N1m

5	2584	96.8	3846	13	ADK52133 Human N1m
6	2531.5	94.8	3128	13	ADK54234 Human N1m
7	2499	93.6	3147	14	ADV97751 Human N1m
8	2409	90.2	3304	4	AAE44681 Human N1m
9	2409	90.2	3304	12	ADJ29379 Human N1m
10	2368	88.7	3538	13	ADV35041 Human N1m
11	2368	88.7	3538	14	AAE41757 Human N1m
12	2301	85.4	1501	4	AAK51824 Human N1m
13	1617	60.6	981	5	AAE29899 Human N1m
14	1534.5	57.5	2141	5	AAH14904 Human N1m
15	1534.5	57.5	2158	6	ABA93734 Human N1m
16	1534.5	57.5	2163	10	ADG14684 Human N1m
17	1534.5	57.5	2163	10	ADG10731 Human N1m
18	1534.5	57.5	2163	12	ADJ29382 Human N1m
19	1534.5	57.5	2163	14	ADZ49553 Human N1m
20	1534.5	57.5	2163	14	AAE41756 Human N1m
21	1534.5	57.5	2225	12	ABK43561 Human N1m
22	1534.5	57.5	2225	12	ADJ53948 Human N1m
23	1529.5	57.3	2159	5	AAH14231 Human N1m
24	1529.5	57.3	2159	5	AAH78063 Human N1m
25	1522.5	57.0	2178	10	ADG10843 Human N1m
26	1508.5	56.5	2241	2	AAV59121 Human N1m
27	1502.5	56.3	1931	3	AAE98922 Human N1m
28	1360.5	51.0	2328	4	AAE44682 Human N1m
29	1360.5	51.0	2328	12	ADJ29380 Human N1m
30	1174.5	44.0	2803	4	ABJ02019 Human N1m
31	1170	43.8	2923	8	ABE80409 Human N1m
32	1106.5	41.4	1568	4	AAE44685 Human N1m
33	1106.5	41.4	1568	12	ADJ29383 Human N1m
34	1087.5	40.7	1606	5	AAE64747 Human N1m
35	1022.5	38.3	5655	4	ABJ02018 Human N1m
36	965.5	36.2	1980	12	ADJ58748 Human N1m
37	955	35.8	1065	4	AAE27150 Human N1m
38	955	35.8	1065	4	ABK43665 Human N1m
39	955	35.8	1065	10	ADJ29382 Human N1m
40	955	35.8	1065	12	ADJ54252 Human N1m
41	810.5	30.4	2002	12	ADJ5884 Human N1m
42	756	28.3	538	10	ADJ22646 Human N1m
43	756	28.3	910	4	AAE27174 Human N1m
44	756	28.3	910	4	ABK43562 Human N1m
45	756	28.3	910	10	ADJ93352 Human N1m

ALIGNMENTS

RESULT 1	ADJ96576	ADJ96576 standard; DNA; 3765 BP.
ID	ADJ96576	
XX	ADJ96576	
AC	ADJ96576	
DT	06-MAY-2004	(first entry)
XX		
DE	Human Nim-A related protein kinase NRBP2 DNA SeqID 33.	
XX		
KW	gene; ds; kinase; human; tyrosine protein kinase; serine/threonine protein kinase; PTK; gene therapy; cancer	
KW	immune-related disease; cardiovascular disease; brain;	
KW	neuronal associated disease; metabolic; inflammatory disorder;	
KW	cytosolic; neuroprotective; immunomodulator; antiinflammatory;	
KW	Nim-A related protein kinase; NRBP2.	
XX		
OS	Homo sapiens.	
OS	53.	
XX		
PN	WO2004006838-A2.	
XX		
PD	22-JAN-2004.	
XX		
PF	15-JUL-2003; 2003WO-US021730.	
XX		
PR	15-JUL-2002; 2002US-0395632P.	
XX		

PA (SUGEN-) SUGEN INC.
XX Whyte D, Manning G, Caenepeel S;
XX WPI: 2004-122753/12.
DR P-PSDB; ADJ96642.
XX
XX New nucleic acid molecule encoding a kinase polypeptide, useful for
PT preparing a composition for treating diseases or disorders, e.g., cancer,
PT or neurological, immunological or inflammatory disorders.
XX
XX Example 1; SEQ ID NO 33; 366pp; English.
XX
XX This invention relates to a novel isolated, enriched or purified nucleic
CC acid molecule that encodes a kinase polypeptide. Specifically, it relates
CC to human tyrosine and serine/threonine protein kinases (PK's and STK's),
CC as well as protein kinase-like enzymes. The present invention describes
CC screening methods to identify agonists, antagonists and antibodies that
CC can be used to modulate the activity or function of the mammalian kinase
CC enzymes. As such, these compositions can be used for gene therapy
CC purposes to treat diseases or disorders including cancer, immune-related
CC diseases, cardiovascular diseases, brain or neuronal associated disease,
CC metabolic and inflammatory disorders. Accordingly, they exhibit
CC cytostatic, neuroprotective, immunomodulator and anti-inflammatory
CC activities. This polynucleotide sequence is a human kinase DNA sequence
CC of the invention.
XX
SQ Sequence 3765 BP; 801 A; 1168 C; 1040 G; 756 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,046-259 Length: 3765
Score: 2670.00 Matches: 507
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 12 Gaps: 0

US-10-618-941-99 (1-507) x ADJ96576 (1-3765)
QY 1 MetAlaAlaProGluProAlaProAlaProAlaArgAlaArgGluArgGluArgGluArg 20
Db 282 ATGGGGGGCCCCCGAGCCCGCGCGCGAGGCGGGCCCGGAAACGGGAGCGGAGCGAGAGC 341
QY 21 GluSerGluAapGluSerAspIleLeuGluGluSerProGlyAlaArgTrpGlnArg 40
Db 342 GAGAGCGAG 401
QY 41 ArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
Db 402 CGGAGCGAGGTAAACCAAGGAGCAATGCCAGGGCTTCAGAGCACTTCCTGAGCCATGAGC 461
QY 61 ThrGluGluGlyValGluValAlaTrpAsnGluLeuAsnIlePheGlyAspArgIleAspHe 80
Db 462 ACGGAGGAGGGGGGTGAGAGGTGTGAGAACGAGCTCCACTTCGAGAGACGAGAGGCCCTTC 521
QY 81 AlaAlaHisGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 100
Db 522 GGGGGCGAG 581
QY 101 AsnIleValIleLeuHisIleValSerGlySerLeuGluGlnPheLeuIleValIle 120
Db 582 AACATCGTAGATTGCAACAAGTACGCTGAGTACCTCTGAGGCTTCGCGAGGGGTGATC 641
QY 121 PheIleThrGluTrpValSerSerGlySerLeuGluGlnPheLeuIleValIle 140
Db 642 TTCATCAGAGAGTACGTGTCATCAGGAGCTCAAGCAATTCCTCAAAAAGCAAGAG 701
QY 141 AsnHisIleValMetAsnAlaArgAlaTrpGlyArgTrpGlnIleLeuSerAla 160
Db 702 AACCAAGAGGCGATGAGAGCCCGGGGCTGAGAGCCGTGTGACACGAGTCTGCTCGG 761
QY 161 LeuSerPheLeuHisAlaCysSerProGlyIleIleHisGlyAsnLeuThrSerAspThr 180

Db 762 CTCAGCTTCCTGCAGCGCTCGAGCCCCCAATCATCCAGGAAACCTGACGAGACACC 821
QY 181 IlePheIleGlnHisAsnGluLeuIleValIleGlySerValIleTrpHisArgIlePheSer 200
Db 822 ATCTTCATTCAGCAACAAGGCTTCATCAAGATCGGCTCGGTGAGCAAGCAATCTTCCTCC 881
QY 201 AsnAlaLeuArgProGlyProThrAlaLeuProAspAspLeuArgSerProIleArgAlaGlu 220
Db 882 AATGCTCTGGCCCTTCACAGCACTTCAGATATCTCCGAAGCCCATTCGGCGCTGAG 941
QY 221 ArgGluGluLeuArgAsnLeuHisPhePheProGluTrpGlyGluValAlaAspGly 240
Db 942 CGAGAGGAACTTCGGAACCTGCACTTCCTCCCAAGATGTGAGAGGTGGCGATGG 1001
QY 241 ThrAlaValAlaPhePheSerPheGlyMetCysValAlaLeuGluMetAlaValLeuGluIle 260
Db 1002 ACCGCTGTGACATCTTCCTTTGGAGATGTGCGGTGAGATGGCTGTACTGGAAATC 1061
QY 261 GlnThrAsnGluAapThrArgValThrGluGluAlaIleAlaArgAlaArgHisSerLeu 280
Db 1062 CAGACCAATGGAGACACCGGGTCAAGAGAGGCCATTCCTCGCGCAAGCACTCGCTG 1121
QY 281 SerAspProAsnMetArgGluPheIleLeuCysCysLeuAlaArgAspProAlaArgArg 300
Db 1122 AGTACCCCAACATGCGGAGATTCATCTTGTGCTGCTGCGCCGGAGACCTGCGCGCGG 1181
QY 301 ProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeuIleLeu 320
Db 1182 CCCCTGCCCCACACCTTCCTTCACCGCGGTGCTTCGAGGTGCATCTGCTAAGGTC 1241
QY 321 LeuAlaAlaHisCysPheIleGlnHisGlnTrpLeuMetProGluAsnValAlaGluGlu 340
Db 1242 CTGGAGCCGCCACTGCTTCATCAGACACAGTACCTCATGCTTGAGATGTGTGAGAGAG 1301
QY 341 LysThrIleValMetAspLeuHisAlaValLeuAlaGluLeuProArgProArgArgPro 360
Db 1302 AAGACCAAGGCGCATGAGACTGCAAGCGGTCTTGCGAGGCTTCCCGGCGCCGAGGCC 1361
QY 361 ProLeuGlnTrpArgGlyTrpSerGluValSerPheMetGluLeuAapIlePheLeuGluAsp 380
Db 1362 CGGTGCGAGTGGCGGTCTCGGAAGTCTCTTCATGTGAGCTGAGCAAAATTCCTGAGGAT 1421
QY 381 ValArgAsnGlyIleTrpProLeuMetAsnPheAlaIleThrArgProLeuGluIlePro 400
Db 1422 GTCCAGAAATGAAATCTTACCACTATGAACTTTGACGACCACTCAACCTCTGGGCTGCC 1481
QY 401 ArgValLeuAlaProProProGluGluValGlnIleValAlaIleValTrpProThrProGluPro 420
Db 1482 CGTGTGCTGGCCCAACCCCGGAGAGGTCCAAAAGGCCAAGACCCGAGCCAGAGGCC 1541
QY 421 PheAspSerGluTrpArgIleValIleGlnMetGlnCysAsnLeuGluArgSerGluAsp 440
Db 1542 TTTGACTCTGAGACAGAGAAAGTATCATCAGATGAGTGCACACTCGAGAGAGAGAGAGC 1601
QY 441 LysAlaIleArgTrpHisIleuThrLeuLeuValLeuGluIleAspArgLeuHisIleArgGlnLeu 460
Db 1602 AAGGCGGCTGCAATCTACTCTCTGCTGTGAGAGACCGGCTGACCGGAGGCTG 1661
QY 461 ThrTrpAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuValHisIleTrp 480
Db 1662 ACCATCAGACCTGCTCCCAAGGAGAGAGGCCCAAGAGACCTCGGCTGAGACTGTGACATAT 1721
QY 481 GlyPheLeuHisGluAapAspArgMetCysLeuAlaIlePheLeuGluSerThrPheLeu 500
Db 1722 GGGTTCCTCCACAG 1781
QY 501 LysTrpArgGlyThrGlnAla 507
Db 1782 AAGTACCGGTGGAGCCAGGCC 1802

RESULT 2
ADJ13324
ID ADJ13324 standard; cDNA; 4048 BP.

Db 1464 CGTGTGCTGGCCCAACCCCGAGAGAGGTCCAAAAGGCCAAGACCCCGGACCGCAGAGCC 1523
Qy 420 OPHeaSPeSergIuThrAglYValIIGImeGInCYsAnLeuGIuArgSergIuAs 440
Db 1524 CTTGACTCTGAGACGAGAAAGGTATCCAGATGACGTGCAACTGGAGAGAGCGAAGA 1583
Qy 440 pLyAlaRgTrpHIsleuThrLeuLeuValIleuGIuAspAglYLeuHIsArgInLe 460
Db 1584 CAAGGCGGCTGGCACTCCTCTCTGCTGTGAGAGACCGGCTGCACCGGCACT 1643
Qy 460 uThrYrAspLeuLeuProThrAspSerAlaGInAspLeuAlaSerGIuLeuValHIsTy 480
Db 1644 GACCTACGACCTGCTCCCAAGGAGAGAGGCGCCAGACCTCGCTCGAGCTCGTGCATTA 1703
Qy 480 rGIYpHeuHIsGIuAspAspArgMetIYsleuAlaAlaPheLeuGIuSerThrPheLe 500
Db 1704 TGGCTTCTCCACGAGAGACGACCGGATGAGCTGCGCTTCTTGAGAGACCTTCTCT 1763
Qy 500 uLyfTrArgIYrGInAla 507
Db 1764 CAAGTACCTGGAGACCGCAGGCC 1785
RESULT 3
ADXS2130
ID ADXS2130 standard; cDNA; 3837 BP.
AC ADXS2130;
XX 21-APR-2005 (first entry)
DE Plant full length insert polynucleotide seqid 26870.
XX plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX unidentified.
OS US2004034888-A1.
PN 19-FEB-2004.
PD 28-APR-2003; 2003US-00425114.
PF 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX 06-MAY-1999; 99US-00304517.
PR 05-NOV-2001; 2001US-00985678.
XX (LIUJ/) LIU J.
PA (ZHOU/) ZHOU Y.
PA (KOVA/) KOVALIC D K.
PA (SCRE/) SCREIB S E.
PA (TABA/) TABASKA J E.
PA (CAOY/) CAO Y.
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX WPI; 2004-180133/17.
XX New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX Claim 1; SEQ ID NO 26870; 15pp; English.
XX The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence

CC available in electronic form from the US patent office at
CC ftp.segdata.uspo.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX SQ Sequence 3837 BP; 799 A; 1195 C; 1084 G; 759 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 2,099-250 Length: 3837
Score: 2584.00 Matches: 501
Percent Similarity: 88.05% Conservative: 0
Best Local Similarity: 88.05% Mismatches: 0
Query Match: 96.78% Indels: 68
DB: 13 Gaps: 2
US-10-618-941-99 (1-507) x ADXS2130 (1-3837)
Qy 1 MetAlaAProGIuPProAlaProArgArgAlaRGIuAArgIuAArgIuAAsp 20
Db 231 ATGGCGGCGCCCGAGACCGCGCGCGGCGCGGCGGCGGAGCGGAGCGGAGGAGC 290
Qy 21 GluSerGIuAspGIuSerAspIleLeuGIuGIuSerProCYsGIYrGTrpGInLYsArg 40
Db 291 GAGAGCAGAGACGAGACGACATCTCGAGAAAGCCGTrGTGCTCGCTGGCAAAAGCGA 350
Qy 41 ArgGIuInValAngInGIYAsnMetProGIYLeuInSerThrPheLeuAlaMetAsp 60
Db 351 CGGAGGAGGTAAACCAAGGGAACATGCCAGGCTTCAAGACACTTCTTAGCCATGGAC 410
Qy 61 ThrGIuGIuGIYValGIuValIValTrpAsnGIuLeuHIsPheGIYAspArgLYsAlaPhe 80
Db 411 ACGAGAGAGGGGGrMAAGTGTGTGTGAACGAGCTCCACTTCGAGACAGGAAGGCTTC 470
Qy 81 AlaAlaHIsGIuGIYsIleGIuThrValPheGIuInLeuValIAspHIsPro 100
Db 471 GCGGCGCAGAGAGAGATCCAGACCGTGTTCGAGACGCTGGTGTGTGACACCGG 530
Qy 101 AsnIleValIYsLeuHIsLYrTrpLeuAspThrSergIuAlaCYsAlaArgValIle 120
Db 531 AACATCGTGAAGTGCACCAAGTACTGGCTGTGATACCTCTAGAGGCTTCGCGAGGTATC 590
Qy 121 PheIleThrGIuTrpValIserSergIYSerIleuYsGInPheLeuLYsTrpThrLYsLYs 140
Db 591 TTCATTCAGAGTACGTGTCTATCAGGCAAGCTCTCAAGCAATTCCTCAAAAAGACCAAGAG 650
Qy 141 AsnHIsLYsAlaMetAsnAlaArg----- 148
Db 651 AACCAAGGCGCATGAAGCGCCGGGTATGGGAGCGGGCTGGGAGCACAGGGGACAGG 710
Qy 148 ----- 148
Db 711 ACGGGGTTGGGCGAGCCTCGGGGACTGGAGTGTGAGGGGGGTGCCCGGCGGCTCGAGCA 770
Qy 148 ----- 148
Db 771 GGGGCTGGGAGGATCGGGGCGGGCTCCGAGAGGCCAGCGGCTTCTCGGCGCCACC 830
Qy 149 -----AlaTrpLYsArgTrpCYsTrpGInIleu 158
Db 831 CGACCGAGGAGTGTGTGCTCGCGCGGAGGCTCGAAAGGCGTGTGTGACGACATCTCTG 890

```

QY 159 SerAlaLeuSerPheLeuH1saIaCySerProIle1leH1sG1yLeuLeuThrSer 178
DB 891 TCGGGCTCAGCTTCTTGACCGCTGCAGCGCCCAATCATCCAGGAACCTGACACAC 950
QY 179 AapThrIlePheH1leG1nH1saH1g1yLeu1leYle1leG1ySerVal1TrpH1saYg1le 198
DB 951 GACACCATCTTCATTCAGACCAACGCGCTCATCAATGCGGCTCCGCTGCGACCGAATC 1010
QY 199 PheSerAspAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArg 218
DB 1011 TTCTCCAAAT-----GCACTTCAGAAATGATCTCCGAAGCCCATCCCG 1052
QY 219 AlAG1uArgG1uG1uLeuArgLeuLeuH1sPhePheProG1uTrg1yG1uVal1a 238
DB 1053 GCTGAGCGAGAGAACTTCGGAACCTGCACCTTCTCCCGACAGATATGAGAGGTGGCC 1112
QY 239 AapG1yThrAlaVala1leP1lePheSerPheG1yMetCyaAlaLeuG1uMetAlaVal1leu 258
DB 1113 GATGGGACCGCTGTGGACATCTTCTCTTGGGATGTGTGCGCTGAGATGGCTGTACTG 1172
QY 259 Glu1leG1nThrAsnG1yAspThrArgVal1ThrG1uG1uAl1a1leAlaArgAlaArgH1s 278
DB 1173 GAAATTCAGACCAATGGAGGACACCCGCGTCACAGAGAGGCGCATTCCTCGCGCAGGAC 1232
QY 279 SerLeuSerAspProAspMetArgG1uPhe1leLeuCySeLeuAlaArgAspProAla 298
DB 1233 TCGCTGAGTGACCCCAACATGCGGAGTTCATCTTGTGCTGCGCCCGGACCTGTGC 1292
QY 299 ArgArgProSerAlaH1sSerLeuLeuPheH1saYg1Val1leuPheG1uVal1H1sSerLeu 318
DB 1293 CCGCGGCCCTGTGCGCACAGCCCTCTTCCACCGGCTGCTGTGAGGTGACCTGTGCG 1352
QY 319 LysLeuLeuAlaAlaH1sCyPheH1leG1nH1sG1nTrg1leuMetProG1uVal1Val1 338
DB 1353 AAGCTCTGTGGACGCCCACTGCTCATCCAGCACCACTACTGCTCATGATGTGTGTG 1412
QY 339 GluG1uLysThrIleVala1leMetAspLeuH1saIaVal1leuAlaG1uLeuProArgProArg 358
DB 1413 GAGGAGAGAGACCAAGGCCATGACCTGCACGCGGTTTGGCGAGCTTCCCGGCCCC 1472
QY 359 ArgProProLeuG1nTrpArgTrg1ySerG1uVal1SerPheMetG1uLeuAspLysPheLeu 378
DB 1473 AGGCCCGCGCTCAGTGGCGGTACTCGAAGTCTCTTCATGAGAGCTGACAAATTCG 1532
QY 379 GluAspVala1ArgAsnG1y1leTrg1ProLeuMetAspPheAlaAla1ThrArgProLeuG1y 398
DB 1533 GAGGATGTCAAGAAATGAAATCTAACCACTGATGAACCTTGGCAGGCCACTCGACCCCTGGGG 1592
QY 399 LeuProArgVal1leuAlaProProProG1uG1uVal1G1nLysAlaG1yThrProThrPro 418
DB 1593 CTGCCCGGTGTCTGGCCCCCGGAGGAGGTTCMAAAGGCCMAAGCCCGGACGCCA 1652
QY 419 GluProPheAspSerG1uThrArgLysVal1leG1nMetG1nCySeAsnLeuG1uArgSer 438
DB 1653 GAGCCCTTTGACTGTGAGACCAAGAAAGTCACTCCAGATCACTGACCACTGAGAGAAC 1712
QY 439 GluAspLysAlaArgTrpH1sleuThrLeuLeuLeuVal1leuG1uAspArgLysH1sArg 458
DB 1713 GAGGACAAAGCGCGCTGCATCTCACTGTCTGTGTGTCGTGGAAGACCGGCTGACCGG 1772
QY 459 GluLeuThrTrg1yAspLeuLeuProThrArgAspSerAlaG1nAspLeuAlaSerG1uLeuVal1 478
DB 1773 CAGCTGACTACGACCTGTCTCCCAAGGACAGACCGCCGAGACTCGCTCGAGGCTCGG 1832
QY 479 H1sTrg1yG1yPheLeuH1sG1uAspAspArgMetLysLeuAla1PheLeuG1uSerThr 498
DB 1893 CACTATGGCTTCTCTCAAGAGACACACCGGATGAAGCTGCGGCTTCTGAGAGACACC 1892
QY 499 PheLeuLysTrg1yArgG1yThrG1nAla 507
DB 1893 TTCTCTCAAGTACCGTGGACCAAGGCC 1919

```

```

ADK52108
ID ADK52108 standard; cDNA; 3838 BP.
XX
AC ADK52108;
XX
DT 21-APR-2005 (first entry)
XX
DE Plant full length insert polynucleotide seqid 26848.
XX
KW plant protectant; plant growth regulant; gene therapy; plant;
KW recombinant DNA construct; physical array; plant breeding marker;
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
KW extreme osmotic condition; pathogen tolerance; pest tolerance;
KW growth rate; cell cycle pathway; disease resistance;
KW galactomannan production; lignin production; plant growth regulator;
KW yield; plant growth; plant development; seed oil; protein yield;
KW protein content; gene; ss.
XX
OS Unidentified.
XX
EN US2004034888-A1.
XX
PD 19-FEB-2004.
XX
PF 28-APR-2003; 2003US-00425114.
XX
PR 06-MAY-1999; 99US-00304517.
XX
PR 05-NOV-2001; 2001US-00985678.
XX
PA (L1UJ/) LIU J.
XX
PA (ZHOU/) ZHOU Y.
XX
PA (KOVA/) KOVALIC D K.
XX
PA (SCRE/) SCREEN S E.
XX
PA (TABA/) TABASKA J E.
XX
PA (CAOY/) CAO Y.
XX
PI Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
XX
DR WPI; 2004-180133/17.
XX
PT New recombinant DNA construct, useful for improving plant tolerance to
PT cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
PT pests, for conferring increased resistance to plant disease, or for
PT improving yield.
XX
PS Claim 1; SEQ ID NO 26848; 15pp; English.
XX
CC The invention describes a recombinant DNA construct comprising a
CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp.segdata.nepco.gov/sequence.html?docID:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
SQ Sequence 3838 BP; 803 A; 1194 C; 1077 G; 764 T; 0 U; 0 Other;

```

Alignment Scores: 2,09e-250 Length: 3838
Pred. No.: 2584,00 Matches: 501
Percent Similarity: 88.05% Conservative: 0
Best Local Similarity: 88.05% Mismatches: 0

Query Match:	96.78%	Indels:	68
DB:	13	Gaps:	2

QY	1	MetAla1a1aProGluProAlaProArgaArgAlaArgGluArgGluArgGluArgGluArg	20
Db	208	ATGGGGGGCCCCGGACCCGGCCGGAGCGGGCCGGAAACGGGAACGGGAACGGGAAGAC	267
QY	21	GluSerGluAspGluSerAsp11eL euGluGluSerProCyG61YArg1Trop1LnybArg	40
Db	268	GAGACCGAGAGACGAGACGACATCTCGAGAGAAACCCCTGTGTCTCGCTGGCAAAAGCA	327
QY	41	ArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp	60
Db	328	CGGGAGCAGGTAAACCAAGGAGAAACATGCCAGGGCTTCAGAGCACTTCCTAGCCATGAGC	387
QY	61	ThrGluGluGluValGluValVal1TPrpAsnGluuAsn1PheGlyLysAspArgGlyAlaPhe	80
Db	388	ACGGAGAGAGGGGGTGAAGGTGTGTGTGAACGAGCTCCACTTCGGAGACAGAAAGGCTTTC	447
QY	81	AlaAlaH1eGluGluLys1eGlnThrVal1PheGluGlnLeuValLeuValAspH1ePro	100
Db	448	CGGGGCGACGAGAGAAAGATCCAGAACCGTGTTCGAGAGACTGTGTCTGGTGAACACCCG	507
QY	101	Asn11eValLysLeuH1eLys1eTyr1TPrpLeuAspThrSerGluAlaCyseAlaArgVal11e	120
Db	508	AACATCTGTGAAGTTGCACAAAGTACTGCTGGATACCTCGAAGGCTGGCCGAGGGTCAATC	567
QY	121	Phe11eThrGluTyrValSerSerGlySer1eLysLysGlnPheLeuLysLysThrTyrLys	140
Db	568	TTCAATCAAGAGTACTGTCTCATCAAGGACCTCAAGACAAATCTCTCAAAAGCAAGAGAG	627
QY	141	AsnH1eLysAlaMetAsnAlaArg-----	148
Db	628	AACCAACAAGGCCATGAAAGCCCGGGATAGGGGAGCGGGCTGGGACGACAGGGGACAG	687
QY	148	-----	148
Db	688	ACGGGGTTGGGGCAGCCTCGGGAACTGGGATGTGAAGGGAGTCCCGCGGCTCGACA	747
QY	148	-----	148
Db	748	GGGGCTGGGCGAGATCGGGGGCGGCTCCGAGGCCAGCGCGCTCTCGCGCCAC	807
QY	149	-----AlaTPrpLysArgTPrpCyThrGln1Leu	158
Db	808	CGACCGACGAGTCTGTCGCTCCGCCCGCAGCGCTGGAAGCGCTGTGTGACGACGATCTCTG	867
QY	159	SerAlaLeuSerPheLeuH1eAlaCyseSerProProl1e11eH1eGlyAsnLeuH1eSer	178
Db	868	TTTGAGCTAGCTTCTCGACGCTCGACGCCCCCAATATCAAGGAAACCTGTACAGC	927
QY	179	AspThr1ePhe11eGlnH1eAsnGlnLysLeu11eLys11eGlySerVal1TPrH1eAspGly1e	198
Db	928	GACACCATCTTCAATTCAGACAAAGGCTCATCAAGATGGCTTCGTGTGTGAGACGGAATC	987
QY	199	PheSerAsnAlaLeuArgProThrAlaLeuProAspAspLeuArgSerPro11eArg	218
Db	988	TTCTCTCAAT-----GCACCTTCAGATGATCTCCGAAGCCCAATCCGC	1028
QY	219	AlaGluArgGluGluLeuArgAsnLeuH1eAspPheProProlGluTyrGlyValAla	238
Db	1030	GCTGAGCGAGAGAACTTCGAAACCTGTCACTTCTTCCCCCAAGATGAGAAAGGTGGCC	1088
QY	239	AspGlyThrAlaValAsp11ePheSerPheGlyMetCyseAlaLeuGluMetAlaValLeu	258
Db	1090	GATGGGACCGCTGTGGACATCTTCTCTTGGAGTGTGTGGCTGGAGATGGCTGTACTG	1144
QY	259	Glu11eGlnThrAsnGlyAspThrArgVal1ThrGluGluAla11eAlaArgAlaArgH1e	278
Db	1150	GAATATCCAGACCAATGTGGGACACCGGGGTCAAGAGAGGCCATGTCTCGCCCAAGGAC	1208

QY	279	SeriesSerAProAmeMetArgGlubHe1leuCySvSleuA1aArgProA	298
Db	1210	TCGCTGATGTAGCCCCCAATCGGGAGTTCATCTTGTGCTCGGGAGCCCTCC	1265
QY	289	ArgArgProSerA1aH1aSerLeuLeuHe1aArgVal1LeuPheGluValH1aSerLeu	318
Db	1270	CGCGGCCCCCTTGCCACAGCTCTCTTCCACCGCGGCTCTTCAGAGTGCATCGGCTG	1322
QY	319	LySLeuLeuA1aA1aH1aSerPhe11eGlnH1eGlnTyrLeuMetProG1uAaAVal1	338
Db	1330	AAAGCTCCGGGAGGCCCATGCTGTTATCCAGACCAAGTACCTTCATGCTCGAGAAATGGTG	1389
QY	339	GlUGluTyHTrLySAlaMeAspLeuH1a1aVal1aGluLeuProArgProArg	358
Db	1390	GAGGAGAAAGACCAAGCCATGGAACCTGCACCGGCTTGGCGAGCTTCCCGGCGCCGC	1445
QY	359	ArgProProLeuGlnTrpArgTySerGluValSerPheMetGluLeuAspArgPheLeu	378
Db	1450	AGGCCCCCGGTGCAGTGGCGGTACTCGAAAGTCTTCATAGAGACTGACAAATTCCTG	1506
QY	379	GlubAspValaArgAaGly11eTyProLeuMetAsnPheA1aA1aThrArgProLeuGly	398
Db	1510	GAGAGATGTCAAGAAATGAAATTAACCACTGATGAATTTGACAGCACCTCGACCCCTGGGG	1566
QY	399	LeuProArgVal1leuA1aProProProGluGluValGlnTySAlaTyHTrProThrPro	418
Db	1570	CTGCCCCGTGTGCTGGCCCCCAACCCCGGAGAGGTCCAAAGGCCAAGACCCCGACGCA	1622
QY	419	GlUProPheAspSerGluTrpArgLySAl11eGlnMetGlnCyAaLeuGluArgSer	438
Db	1630	GAGCCCTTTGACTGTGAGACCAAGAAAGTCTCCAGAAGTGACAGTGTACCTGAGAGAAAGC	1689
QY	439	GlubAspValaArgTrpH1eLeuThrLeuLeuLeuVal1LeuGluAspArgLeuH1aArg	458
Db	1690	GAGAGCAAGGCGCGCTGCATCTCACTGCTTGTGGTCTGGAGACCGGCTGCACCGG	1744
QY	459	GlnLeuThrTyAspLeuLeuProThrAspSerA1aGlnAspLeuA1aSerGluLeuVal	478
Db	1750	CAGCTGACCTTAGACCTGCTCCCAACGAGACGCGCCAGACCTTGCTCGGAGCTCGTG	1809
QY	479	H1aTyTyGlyPheLeuH1aGluAspAspArgMetLySleuA1aA1aPheLeuGluSerThr	498
Db	1810	CACATATGCTTCTCCCAAGAGACGACCGGATGAAGCTGCTGCTTCTTGAGAGACAC	1865
QY	499	PheLeuLySTyrArgGlyThGlnA1a507	
Db	1870	TTCTCTCAAGTACCGTGGAGCCAGGCC1896	
RESULT 5			
ADX52133			
ID	ADX52133 standard; cDNA; 3846 BP.		
XX	ADX52133;		
DT	21-APR-2005 (first entry)		
XX	Plant full length insert polynucleotide seqid 26873.		
DE	plant proteccant; plant growth regulant; gene therapy; plant;		
KM	recombinant DNA construct; physical array; plant breeding marker;		
KM	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;		
KM	extreme osmotic conditioin; pathogen tolerance; pest tolerance;		
KM	growth rate; cell cycle pathway; disease resistance;		
KM	galactomannan production; lignin production; plant growth regulator;		
KM	yield; plant growth; plant development; seed oil; protein yield;		
XX	protein content; gene; ss.		
XX	Unidentified.		
XX	US2004034888-A1.		
PN	19-FEB-2004.		
PD			
XX			

[illegible]

OY	101	AsnIleValIlyLeuHniSlyeTyTrIrpLeuAspTrpSerGluAlaCysAlaArgValIle	120
Db	516	AACATCGTAAGTTCACAAAGTACTGGCTGGATACCTCTGAAGCGTCGCGAAGGCATC	575
OY	121	PheIleTrnGluTyValSerSerGlySerLeuLysGlnPheLeuIlyblyThryIlys	140
Db	576	TTTCATCAGAGTACTGTGCATCAGGCAAGCCTCAAGCAATTCTCTCAAAAAGCAAGAG	635
OY	141	AsnHilSlybAlaMetAsnAlaArg-----	148
Db	636	AAACCAAGGCGCATGAACGCCGGGTATATGGAGACGGGCTGGGGCAGCCACGGGACAGG	695
OY	148	-----	148
Db	696	ACGGGGTTGGGGCAGCCTTCGGGAACTGGAGTGTGAAGGGGTGTCCCGCGGCTCGAGCA	755
OY	148	-----	148
Db	756	GGGGCTGGGCGAAGATGCGGGGCGGGCTCCGAGGCCAGCGCGCTCTCTGGGCCACC	815
OY	149	-----AlaTrpLysArgTrpCysThrglnIleLeu	158
Db	816	CGACCGACGAGTCTGGTGCTCCGCCGCCAGCGCTGGAAGCGTGGTGCAACGAGATCTCG	875
OY	159	SerAlaLeuSerPheLeuHniSAlaCysSerProProIleIhIhIglYanLeuTrsSer	178
Db	876	TCTGGGCTCAGGTTCTCTCAGCCCTGCAGCCCCCAATCATCACGGGAACCTGCACAG	935
OY	179	AspThrIlePheIleGlnHniSasnGlyLeuIlySileGlySerValTrpHisArgIle	198
Db	936	GACACCATCTTCATTACAGCACACGGCTCATCAAGATCGGCTCGTGGACCGCAATC	995
OY	199	PheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArg	218
Db	996	TTCTCCAT-----GCATTCAGATGATCTCGAAGCCCATCCGC	103
OY	219	AlaGluArgGluGlnLeuArgAsnLeuHniSphePheProProGluTyTrGlyValAla	238
Db	1038	GCTGAGCAGAGGAACCTTCGGAACTGCACCTTCCCCCGAGATATGAAAGGTGGCC	109
OY	239	AspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeu	258
Db	1098	GATGGAGCCGTGTGGACATCTTCTCTTGGAGTGTGGCCCTGAGATGGCTGTACG	115
OY	259	GluIleGlnThrAsnGlyAspThrArgValThrGluGlnAlaIleAlaArgAlaArgHis	278
Db	1158	GAATTCAGACCAATGGGAGACCCGGGTACAGAGAGGCCATTGCTTGGCCCAAGGCAC	121
OY	279	SerLeuSerAspProAsnMetArgGluPheIleLeuCysCysLeuAlaArgAspProIa	298
Db	1218	TTCGTGATGATACCCCAATCATGGGGAGTTTACTTGTGTGCTGGCCCGGACCTTGCC	127
OY	299	ArgArgProSerAlaHisSerLeuLeuPheHniSArgValLeuPheGluValHisSerLeu	318
Db	1278	CGCCGGCCCTTCGCCCAAGCCTCTCTTCCACCCGCGTCTTGGAGGTGACCTGCGTG	133
OY	319	LysIleLeuAlaAlaHisCysPheIleGlnHniSglnTyTrLeuMetProGluAsnValVal	338
Db	1338	AAGCTCCGCGAGCCACTGTCTTATCCAGCACCCAGTACTCTATGCTCCTGAGATGTGGTG	139
OY	339	GluGluTyThrIlybAlaMetAspLeuHniSAlaValIleuAlaGluLeuProArgProArg	358
Db	1398	GAGGAGAAAGACCAAGCCATGGAAGCTGTGACCGGTCCTTGGCGAGCTTCCCGGCCGCC	145
OY	359	ArgProProLeuGlnTrpArgTyTrSerGluValSerPheMetGluLeuAspLysPheLeu	378
Db	1458	AGGCCCCCGTGCAGTGGCGGTACTCGGAAGTCTCTTCATGTGAGACTGGACAAATTCTG	151
OY	379	GluAspValaArgAsnGlyIleTyTrProIleuMetAsnPheAlaIleThrArgProIleuGly	398
Db	1518	GAGATGTCAAGAAATGAAATCTTACCACTGTAGAACTTGTGACCGCACTGACCCCTGGGG	157

QY 399 LeuProArGValIleuAlaProProProGluGluValGlnIlybAlaIyTherProthPro 418
DB 1578 CTGCCCCGTGTCTGGCCCCACCCCGAGAGGTCTCAAAAGGCCAAAGCCCCGACGCCA 1637
QY 419 GluProPheAepSerGluThrArgIlyValIleGlnMetGlnCyAsnIleuGluArgSer 438
DB 1638 GAGCCCTTGAAGCTGAGACCGAAAGGTTCATCGAGATGCACTGGAGAGAAAGC 1697
QY 439 GluArapIyAlaArgTrpHisIleuThrIleuLeuValIleuGluAspArgIleuHisArg 458
DB 1698 GAGGACAAAGGCGCGGTGGATCTCATCTGCTTCTGGCTGGAGACCGGCTGCACCGG 1757
QY 459 GlnLeuThrTrpAspIleuLeuProThrAspSerAlaGlnAspIleuAlaSerGluLeuVal 478
DB 1758 CAGCTGACCTTACAGCTGCTCTCCCAACGACAGCGCCGACCTCGAGCTCGTG 1817
QY 479 HisTrGlyIyPheLeuHisGluAspAspArgMetIyIleuAlaAlaPheLeuGluSerThr 498
DB 1818 CACTATGGCTTCTCCACAGAGACGACCGGATGAAGCTGGCCGCTTCTGAGAGACAC 1877
QY 499 PheLeuIyTrpArgGlyThrGlnAla 507
DB 1878 TTCCCTCAAGTACCGTGGAGCCAGGCC 1904
RESULT 6
ADK54234
ID ADK54234 standard; cDNA; 3128 BP.
XX
XX ADK54234;
XX
XX AC
XX
XX DT 21-APR-2005 (first entry)
XX
XX DE Plant full length insert polynucleotide seqid 28974.
XX
XX plant protectant; plant growth regulant; gene therapy; plant;
XX recombinant DNA construct; physical array; plant breeding marker;
XX cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
XX extreme osmotic condition; pathogen tolerance; pest tolerance;
XX growth rate; cell cycle pathway; disease resistance;
XX galactomannan production; lignin production; plant growth regulator;
XX yield; plant growth; plant development; seed oil; protein yield;
XX protein content; gene; ss.
XX
XX OS Unidentified.
XX
XX PN US2004034888-A1.
XX
XX PD 19-FEB-2004.
XX
XX PF 28-APR-2003; 2003US-00425114.
XX
XX PR 06-MAY-1999; 99US-00304517.
XX PR 05-NOV-2001; 2001US-00985678.
XX
XX PA (LNUJ/) LNU J.
XX PA (ZHOU/) ZHOU Y.
XX PA (KOVA/) KOVALLIC D K.
XX PA (SCRE/) SCREEN S E.
XX PA (TABAS/) TABASKA J E.
XX PA (CAOY/) CAO Y.
XX
XX LIU J, Zhou Y, Kovallic DK, Screen SE, Tabaska JE, Cao Y,
XX
XX WPI; 2004-180133/17.
XX
XX New recombinant DNA construct, useful for improving plant tolerance to
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
XX pests, for conferring increased resistance to plant disease, or for
XX improving yield.
XX
XX Claim 1; SEQ ID NO 28974; 15bp; English.
XX
XX The invention describes a recombinant DNA construct comprising a

CC polynucleotide consisting of a sequence encoding an amino acid sequence
CC available in electronic form from the US patent office at
CC ftp:seedata.uspto.gov/sequence.html?docid:2004034888. The polynucleotide
CC of the invention are also useful in physical arrays of molecules and as
CC plant breeding markers. The recombinant DNA construct is useful for
CC improving plant tolerance to cold, heat, drought, herbicides, extreme
CC osmotic conditions, pathogens or pests, for manipulating growth rate in
CC plant cells by modification of the cell cycle pathway, for conferring
CC increased resistance to plant disease, for producing galactomannan,
CC lignin or plant growth regulators, for increasing the rate of homologous
CC recombination in plants, for improving yield by modification of
CC photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
CC or by providing improved plant growth and development under at least one
CC stress condition or for modifying seed oil or protein yield and/or
CC content. This sequence represents a plant full length insert
CC polynucleotide that can be used in the recombinant DNA construct of the
CC invention.
XX
XX Sequence 3128 BP; 617 A; 973 C; 966 G; 572 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3,17e-245 Length: 3128
Score: 2531.50 Matches: 499
Percent Similarity: 83.47% Conservative: 1
Best Local Similarity: 83.31% Mismatches: 1
Query Match: 94.81% Indels: 99
DB: 13 Gaps: 3
US-10-618-941-99 (1-507) x ADK54234 (1-3128)
QY 1 MetAlaIaPProGluProAlaProArGValIleuAlaArgGluArgGluArgGluAsp 20
DB 120 ATGGCGGCGCCCGAGCGCGCGCCGAGCGCGCGCGGAAACGAGCGGAGCGGAGGAC 179
QY 21 GluSerGluAspGluSerAspIleLeuGluGluSerProCyGlyArgTrpGlnIlybArg 40
DB 180 GAGAGCGAGACCGAGACGACATCTTGGAGAAAGCCGTTGCTCGTGGCAAAAGCGA 239
QY 41 ArgGluGlnValaAngIyAsnMetProGlyLeuGlnSerThrPheLeuAlaMetAsp 60
DB 240 CGGAGACAGTAAACCAAGGAAATGATCCAGAGGCTTCAGACACTTCTAGCATGAGAC 299
QY 61 ThrGluGluGlyValGluValaValaTrpAsnGluLeuHisPheGlyAspArgIlybAlaIle 80
DB 300 ACGAGAGAGGAGGTGAAGGTGTGTGAACGAGCTTCCATTCGAGACAGAGAGGCTTC 359
QY 81 AlaAlaHisGluGluIyIleGlnThrValPheGluGlnIleuValIleuValAspHisPro 100
DB 360 GCGGCGCACAGAGAAATCCACACCGTGTTCAGACAGCTGTGTGTGTGACACACCGG 419
QY 101 AsnIleValIyIleuHisIlybTrpIleuAspThrSerGluAlaCybAlaArgValIle 120
DB 420 AACATCGTGAAGTTCACAAAGTACTGGCTGTGATCTCTGAGGCTCGGAGGGTCAATC 479
QY 121 PheIleThrGluTrpValSerSerGlySerIleuIyGlnPheLeuIyIleuThyIlyb 140
DB 480 TTCATCAACAGTACGTGTCTATCGAGGACCTCAAGCAATTCCTCAAAAAGCAAGAG 539
QY 141 AsnHisIyAlaMetAlaAlaArg----- 148
DB 540 AACCAAAAGCCATGAACGCCCGGATGAGGAGCGGCTGGGCAAGCCACGGGACAGG 599
QY 148 ----- 148
DB 600 ACGGAGTTGGGGCAGCCTCGGGGACCTGGAGATGAGAGGGGTGCCCGGCGCTCGAGCA 659
QY 148 ----- 148
DB 660 GGGGCTGGGCGAGGATCGGGGCGGGCTCGGACAGGCCAGCGCTCTCGGCGCCAC 719
QY 149 -----AlaTrpIyAspArgTrpCybThrGlnIleu 158
DB 720 CGACCGACGAGTGTGTGCTCCGCGCCAGGCTGAGAGCGCTGTGTGACGACGATCTGT 779

QY 159 SerAlaLeuSerPheLeuH1sAlaCySerProProlleH1sGlyAsnLeuThrSer 178
 DB 780 TCTGGCTAGCTTCTTGCAGCGCTGACGCCCCCAATCATCCAGGAACTGACACG 839
 QY 179 AapThrIlePheH1sGlnH1sAangIyLeuIleIyLeuIleGlySerValTrpH1sArgIle 198
 DB 840 GACACCATCTTCTTATGAGCAACGCGCTCATCAATAGTGGCTCGGTGGACCGAATC 899
 QY 199 PheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProlleArg 218
 DB 900 TTCTCCAAAT-----GCACCTTCAGATGATCTCCGAAAGCCCATCCCG 941
 QY 219 AlAGlAArgGluGluLeuArgAsnLeuH1sPhePheProProGluTrpGlyValAla 238
 DB 942 GCTGACGAGAGAACTTGGAACTTGCACCTTCTCCCGCAAGATAGAGAGTGGCC 1001
 QY 239 AapGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeu 258
 DB 1002 GATGGGACCGCTGTGACATCTTCTCTTGGGATGTGTGGCTGGAGATGGCTGTACTG 1061
 QY 259 GluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHis 278
 DB 1062 GAAATTCAGACCAATGGGACACCCCGGTCCACAGAGAGGCCATTGCTCGCCGACGGCAC 1121
 QY 279 SerLeuSerAspProAsnMetArgGluPheH1sLeuCySerLeuAlaArgAspProAla 298
 DB 1122 TCGCTGAGTGAACCCCAACATGGCGAGGTTCATCTTGTGCTGCTGGCCCGGACCTTGCC 1181
 QY 299 ArgArgProSerAlaH1sSerLeuLeuPheH1sArgValLeuPheGluValH1sSerLeu 318
 DB 1182 CGCCGGCCCTGTGCCACAGCCCTCTTCCACCGGTGCTCTCGAGGTGACCTGCTG 1241
 QY 319 LysLeuLeuAlaAlaH1sCySerPheH1sGlnH1sGlnTrpLeuMetProGluAsnVal 338
 DB 1242 AAGCTCTGTGAGCCCACTGCTTCAATCCAGACCACTACCTAGCTGAGATGTGTG 1301
 QY 339 GluGluValThrLysAlaMetAspLeuH1sAlaValLeuAlaGluLeuProArgProArg 358
 DB 1302 GAGGAGAAACCAAGGCAATGACCTGACGCGGTCTTGGCGGAGCTTCCCGGCCCGCC 1361
 QY 359 ArgProProLeuGlnTrpArgTrpSerGluValAspPheMetGluLeuAspLysPheLeu 378
 DB 1362 AGGCCCCCGCTGACAGTGGCGGTACTGCAAGTCTCTTATGAGGTGACAAATTCCTG 1421
 QY 379 GluAspValArgAsnGlyIleTrpProLeuMetAsnPheAlaAlaThrArgProLeuGly 398
 DB 1422 GAGGATGTCAAGAAATGAAATCAACCACTGATGAATTTTGACAGCCACTGACCCCTGGGG 1481
 QY 399 LeuProArgValLeuAlaProProProGluGluValGlnLysAlaLysThrProThrPro 418
 DB 1482 CTGCCCCGTGTGCTGCCCCCAGGAGGTCCAAAGGCCCAAGCCCGGACGCA 1541
 QY 419 GluProPheAspSerGluThrArgLysValIleGlnMetGlnCysAsnLeuGluArgSer 438
 DB 1542 GAGCCCTTGAATCTGAGACCAAGAAAGTCAATCCAGATCAAGTCAACTGAGAGAAC 1601
 QY 439 GluAspLysAlaArgTrpH1sLeuThrLeuLeuValLeuGluAspArgLysHisArg 458
 DB 1602 GAGGACAAAGCGCGTGCATCTCACTGTCTGTGCTGAGAGACCGGCTGACCGG 1661
 QY 459 GlnLeuThrTrpAspLeuLeuProThrAspSerAlaGlnAspLeuAlaSerGluLeuVal 478
 DB 1662 CAGCTGACTTACGACTGTCTCCCAAGCAGCGCCGAGAACCTGCTCGAGTCTGTG 1721
 QY 479 HisTrpGlyPheLeuH1sGlu----- 485
 DB 1722 CACTATGCTTCTTCCACGA-GGTGGCTGGGCGGTGGCGCGGCTTGCGGAGGCGGC 1780
 QY 486 -----AspAspArg 488
 DB 1781 GCACGGGGGACGGCGCCCTCCGATCCCGCATGCTCCTCTCCGAGAGACGACGGG 1840

QY 489 MetLysLeuAlaAlaPheLeuGlnSerThrPheLeuLysTrpArgGlyThrGlnAla 507
 DB 1841 ATGAAGCTGGCGGCTTCTTGGAGAGACACTTCTCAAGTACAGTGGAGCCAGGCC 1897
 RESULT 7
 ID ADV97751
 XX ADV97751 standard; cDNA; 3147 BP.
 AC ADV97751;
 XX
 DT 24-MAR-2005 (first entry)
 XX
 DE cDNA sequence encoding a murine protein kinase Seq 71.
 XX
 KW gene; ser; protein kinase modulator; bioinformatics; diagnostic; cancer;
 KW central nervous system disease; psychotic disorder;
 KW neurological disorder; neurodegenerative disease; metabolic disorder;
 KW cardiovascular disease; inflammatory disorder; gene therapy; cytostatic;
 KW antidiabetic; analgesic; endocrine-gen.; nootropic; tranquilizer;
 KW hypotensive; hypertensive; neuroprotective; antiparkinsonian; vitruide;
 KW fungicide; antibacterial; antidiabetic; anorectic; antiarteriosclerotic;
 KW ophthalmological; antiinflammatory; antiarthritic; antineumatic;
 KW antiasthmatic; osteopathic; antiporiatic; immunosuppressive;
 KW cardiovascular-gen.; vasotropic; antiallergic; gastrointestinal-gen.;
 KW cns-gen.
 XX
 OS Mus sp.
 XX
 PN WO200500200-A2.
 XX
 PD 06-JAN-2005.
 XX
 PF 07-MAY-2004; 2004WO-US014421.
 XX
 PR 09-MAY-2003; 2003US-0469014P.
 XX
 PA (SUBE-) SUGEN INC.
 XX
 PI Caenepeel S, Manning G, Charyczak G, Grigoriev I;
 XX
 DR WPI; 2005-06508/07.
 XX
 P-PSDB; ADV97865.
 XX
 PT New isolated, enriched, or purified kinase nucleic acids and
 PT polypeptides, useful for diagnosing or treating, e.g. cancers,
 PT neurological and neurodegenerative diseases, cardiovascular disease, or
 PT inflammatory disorders.
 XX
 PS Claim 10; SEQ ID NO 71; 300bp; English.
 XX
 CC This invention relates to novel isolated, enriched or purified nucleic
 CC acid molecules that encode kinase polypeptides. Specifically, it refers
 CC to a bioinformatics strategy used to identify mammalian members of the
 CC protein and lipid kinase families. The present invention provides methods
 CC for identifying a substance that modulates the activity of a kinase
 CC polypeptide, as well as a method for the detection of a kinase nucleic
 CC acid in a sample as a diagnostic tool for a disease or disorder.
 CC Furthermore, it describes generation of a knock-out mouse whose genome is
 CC disrupted by recombination at a nucleic acid sequence such that it
 CC produces a phenotype, relative to the wild-type, that exhibits an absence
 CC of kinase activity. The nucleic acids and polypeptides given in the
 CC specification are useful for the diagnosis and treatment of cancer;
 CC central or peripheral nervous system diseases, psychotic and neurological
 CC disorders, neurodegenerative diseases, metabolic disorders,
 CC cardiovascular disease or inflammatory disorders. As such, they can be
 CC used for gene therapy purposes and compositions exhibit cytostatic,
 CC antimigraine, analgesic, endocrine-gen., nootropic, tranquilizer,
 CC hypotensive, hypertensive, neuroprotective, antiparkinsonian, vitruide,
 CC fungicide, antibacterial, antidiabetic, anorectic, antiarteriosclerotic,
 CC ophthalmological, antiinflammatory, antiarthritic, antineumatic,
 CC antiasthmatic, osteopathic, antiporiatic, immunosuppressive, cardiac-
 CC gen, vasotropic, antiallergic and gastrointestinal-gen. This
 CC polynucleotide is a cDNA sequence that encodes a murine protein kinase of

DE Human MARK3-associated cDNA #45.
XX Human; ss; antisense gene therapy; MARK3;
KW MAP/microtubule affinity-regulating kinase 3; cancer;
KW Alzheimer's disease; neurodegenerative disorder;
KW hyperproliferative disorder; cytostatic.
XX Homo sapiens.
XX US2003232771-A1.
XX
XX
XX 18-DEC-2003.
XX
XX 17-JUN-2002; 2002US-00174319.
XX
XX 17-JUN-2002; 2002US-00174319.
XX
XX (ISIS-) ISIS PHARM INC.
XX
XX Ward DT, Freiler SM, Dobie KW;
XX WPI; 2004-052188/05.
XX P-PSDB; ADI29261.
XX
XX
XX New antisense compound targeted to a nucleic acid molecule encoding
PT microtubule-affinity-regulating kinases (MARK3), useful for modulating
PT expression of MARK3 or for treating cancer or Alzheimer's disease.
XX
XX Disclosure; Fig 2; 233pp; English.
XX
XX The invention relates to a compound comprising a sequence comprising 8-80
CC base pairs (bp) targeted to a nucleic acid encoding MARK3
CC (MAP/microtubule affinity-regulating kinase 3), that specifically
CC hybridizes with the nucleic acid encoding MARK3 and inhibits expression
CC of MARK3, i.e., is an antisense oligonucleotide (AO). Also included are a
CC composition comprising the compound and a carrier or diluent, inhibiting
CC the expression of MARK3 in cells or tissues, treating an animal having or
CC suspected of having a disease or condition associated with MARK3 and
CC screening for an antisense compound. The antisense oligonucleotide is
CC useful for preparing a composition for treating hyperproliferative
CC disorder, particularly cancer and neurodegenerative diseases e.g.,
CC Alzheimer's disease. The present sequence is a MARK3 associated cDNA
CC included in the figures but not mentioned anywhere else in the
CC specification.
XX
XX
XX Sequence 3304 BP; 736 A; 1019 C; 841 G; 708 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 8,556-233 Length: 3304
Score: 2409.00 Matches: 460
Percent Similarity: 98.72% Conservative: 2
Best Local Similarity: 98.29% Mismatches: 0
Query Match: 90.22% Indels: 6
DB: 12 Gaps: 1
US-10-618-941-99 (1-507) x ADI29379 (1-3304)
QY 40 Argatggtguglunvalaanglmglyaaemctprogluylenglnserthrphelaualanet 59
Db 1 CAGAGAGAGAAAGGTAAACCAAGGAAACATGCGAGGGCTTCCAGAGCACCTTCCAGCCATG 60
QY 60 Asprrhrgluglugluyvalgluvalvalttrpaenglulwuhiaephelgylaspargluyala 79
Db 61 GACACGAGAGAGGGGGGTAGAGGTGTGTGTGAACGAGCTCCACTTCGAGACAGGAAGGCC 120
QY 80 Phealaalanhieglugluyelleglnthrvalphegluglnleuvalleuvalleuasp 99
Db 121 TTCGGGGGCGACGAGAGAAATCCAGACCGTGTTCGAGCAGCTGTGTGTGTGACAC 180
QY 100 ProaenilevallyaleuuhilvaytrtrpleuaaprrhsersglualaCyaaiaargval 119
Db 181 CCGAACATCTGTAAGTTCACAAAGTAAGTGGCTGATACCTCTGAGGCGCTGGCGAGGCTC 240

QY 120 Ilephellerrhrglugluytrvalserserglvserleuylsglnpheluvalyleuthrllys 139
Db 241 ATCTTCATCAGAGGTACGTCTATCATGACGCGCTCAAGCAATTCCTCAAAAGCCAAAG 300
QY 140 LysaenihslvsaiaametasnaiaargalatrplysargtrpCystrhnglnleuser 159
Db 301 AAGAACCAAGGAGCATGAAGCGCCGGGCTCGAAGCGCTGTGTACACGCAATCTGTCT 360
QY 160 AlauserpheeluuhialaCyaserprobroilellehiesglyaenleuthserasp 179
Db 361 GCGCTCAGCTTCGTCACGCGCTGAGCGCCCATCATCATCCAGGAACTTCACACGCGAC 420
QY 180 ThrlllephellieglinhiaenglyleuileyllegliserValttrphlaargllphe 199
Db 421 ACCATCTTATTCAGCAACAGGCGCTCATGAAGTCGCTCGTGTGACCAATCTTC 480
QY 200 Serasnaialeuargprothrothralaleuproaspaspleuargserprollaargla 219
Db 481 TCCAAAT-----GCACCTTCAGATGATCTCGAAGGCCCATCCGCGCT 522
QY 220 Gluargglugluleuargaeluuhiaephheproprogluytrgluygluvalaasp 239
Db 523 GAGCGAGAGGAACCTCGGAACCTGCACTTCTCCCGCAGAGTATGAGAGGTGGCGCAT 582
QY 240 GlythralvalaaplliepheserPheglwyetCyaaialeuylumetAlaValleuglu 259
Db 583 GGGAGCGGTGGACATCTTCTTGGAGTGTGTGCTGGAGATGGCTGTGTA 642
QY 260 Ileglnthraenglyaaerthrargvalthrugluualaleaiaarglaarghisser 279
Db 643 ATCCAGACCAATGGGAGACCCGGGTCAAGAGAGGCCAATGCTCGGCGCAGGACATCG 702
QY 280 LeuserapProaenmetarglupheilleucyCyaleuualargaspproallaarg 299
Db 703 CTGAGTACCCCAACAGCGGAGTTCATCTTTGCTGCGCGCGGAGACCTGCGCGC 762
QY 300 ArgproserlaahisserleuPhehialargvalleuphegluvalhisserleuyls 319
Db 763 CCGCGCTCTGCGCACGCTCTCTTCCACCGCGCTCTTCGAGGTGCATCTCGCTGAAG 822
QY 320 LeuenualaalanhieCyphellieglinhieglintyrleumectprogluaenvalaigu 339
Db 823 CTCCTGGAGCGCCACTGCTTCATCAGCACCAAGTACTCATGCTGAGAAATGTGTGAG 882
QY 340 Gluylstrhrysaiaametaspleuuhialaaleuualagluuelproatrgproargarg 359
Db 883 GAGAAAGACCAAGGCCATGAGCTTCACGCGGTCTTGCGGAGTCTCCCGCGCGCAGG 942
QY 360 Proproleuglntprargtyrsergluvaliserphemetgluueaplyvphleuglu 379
Db 943 CCCCCGCTGCGAGTGGCTACTCGGAAGTCTCTTCATGAGGTGACAAATTCCTTGAG 1002
QY 380 AspvalarganglyiletyrProleumectasnphealaalatrargproleuglyleu 399
Db 1003 GATGTACGAATGGAATCTAACCTGATGAACCTTTCAGAGCAGCATCAGCCCTGGGCTG 1062
QY 400 ProargvalleualaaproproproglugluvalgluylvalalythrprothrProglu 419
Db 1063 CCCCCGTGTGCGCCCAACCCCGAGAGAGGTCCAAAGGCCCAAGACCCCAAGCCCAAG 1122
QY 420 ProheapserglutrrarglyvalilleglmetglnCyaaenleuGluaargserglu 439
Db 1123 CCGTTGACTCTGAGACCAAGAAAGTATCCAGATGCGATGCAACCTGGAAGAAAGCGAG 1182
QY 440 AsplysaiaargtrphihleuthrleuileuvalleuGluaasparglueuhiaargln 459
Db 1183 GACAAAGGCGGCTGCACTCTACTCTGCTGTGTGTGGAAGACCGGCTGACCGGAG 1242
QY 460 LeuthrtyrApheleuelprothrAspseralaglaapleuualasergluuvalhis 479
Db 1243 CTGACCTACGACCTGTCCCAACGGAAGGCGCCCAAGACCTCGCTCGAGAGCTGTGAC 1302
QY 480 TyrgllyPheleuuhisgluueaprapargmetlybeuualalaapheleugluserThrPhe 499

```

Db      1303 TATGCTTCTCCACAGAGACGCGATGAGCTGCCGCTTCTTGAGAGCACTTC 1362
QY      500 LeuLYTYrArgGLyThrGlnAla 507
Db      1363 CTCAGTACCGTGGAGCCAGGCC 1386

RESULT 10
ADV35041
ID      ADV35041 standard; cDNA; 3538 BP.
XX
XX      ADV35041;
XX
XX      10-FEB-2005 (first entry)
XX
XX      cDNA differentially expressed in the presence of valproate SeqID117.
XX
XX      valproate; 89; multi-parameter high throughput screening; MPHTS;
XX      disease signature; neuropsychiatric; neurodegenerative; schizophrenia;
XX      bipolar affective disorder; BAD; autism; Parkinson's;
XX      Alzheimer's disease; neuroleptic; nootropic; antiepileptic; antidepressant.
XX      Unidentified.
XX      OS
XX      PN      US2003096264-A1.
XX      PD      22-MAY-2003.
XX
XX      18-JUN-2002; 2002US-00175523.
XX
XX      18-JUN-2001; 2001US-0299151P.
XX      PR      07-SEP-2001; 2001US-0317828P.
XX      PR      25-SEP-2001; 2001US-0325150P.
XX      PR      14-NOV-2001; 2001US-033047P.
XX      PR      18-JAN-2002; 2002US-034936P.
XX      PR      04-MAR-2002; 2002US-0361834P.
XX
XX      (PSC-) PSYCHIATRIC GENOMICS INC.
XX
XX      Alter CA, Brockman JA, Evans D, Hook D, Klimczak LJ, Laeng P;
XX      Palfreyman M, Rajan P;
XX
XX      MPI; 2004-118903/12.
XX
XX      Identifying a compound that can treat disease or disorders, such as, a
XX      neuropsychiatric disorder e.g., schizophrenia, or autism, comprises
XX      determining the expression of one or more efficacy genes in a cell
XX      contacted with the test compound.
XX
XX      Claim 9; SEQ ID NO 117; 39pp; English.
XX
XX      This invention relates to a novel screening method identified as a multi-
XX      parameter high throughput screening (MPHTS) assay. Specifically, it
XX      refers to an assay that utilizes the disease signature of a plurality of
XX      specific genes associated with a particular disease, and identifies
XX      differential expression between those cells taken from individuals
XX      affected by that disease and those that are not affected. The present
XX      invention then describes the screening of candidate pharmaceutical
XX      compounds to identify those that have a potential therapeutic benefit for
XX      the treatment of neuropsychiatric and neurodegenerative disorders
XX      including schizophrenia, bipolar affective disorder (BAD) and autism, as
XX      well as Parkinson's and Alzheimer's disease. Accordingly, the compounds
XX      of this invention exhibit various activities including neuroleptic,
XX      nootropic, antiepileptic and antidepressant. Furthermore, the screening
XX      method used in MPHTS will be automated, such that a large number of test
XX      compounds may be rapidly screened with a minimal amount of labour and
XX      effort. This polynucleotide is the cDNA sequence of a gene that is
XX      differentially expressed in the presence of the therapeutic compound
XX      valproate, given in an exemplification of the invention.
XX
XX      Sequence 3538 BP; 806 A; 1076 C; 928 G; 728 T; 0 U; 0 Other;
XX
Alignment Scores:

```

```

Pred. No.: 1,33e-228
Score: 2368.00
Percent Similarity: 87.17%
Best Local Similarity: 86.79%
Query Match: 88.69%
DB: 13
Gaps: 2

US-10-618-941-99 (1-507) x ADV35041 (1-3538)

QY      40 ArgArGGLuGlnValAsnGlnIysAnetProGlyLeuGlnSerThrPheLeuAlaMet 59
Db      1 CAGAGAGAGAAAGTAAACCAAGGAAACATGACGAGGCTTCAGAGCACTTCTCAGCCATG 60
QY      60 AepThrGluGluGlyValGluValValTTPaangIuLeuHisPheGlyAspArgGlyAla 79
Db      61 GACACGAGAGAGAGGGGTAGAGGTGTGTGTGAACGAGCTCCACTTCGAGACAGAGAGGCC 120
QY      80 PheAlaAlaHisGluGluValValTTPaangIuLeuValPheGluGlnLeuValLeuValAspHis 99
Db      121 TTCGCGGCGCAGAGAGAGAGATCCAGACCGTGTTCAGACAGCTGTGTGTGTGTGTGTGTGT 180
QY      100 ProAsnIleValIysLeuHisIlySTyTrPLeuAspThrSerGluAlaCysAlaArgVal 119
Db      181 CCGAACATCTGTAAGTTGACAAAGTACTGTGCTGATACCTCTGAGGCTCGCCGAGGGTCT 240
QY      120 IlePheIleThrGlnIyValSerSerGlySerLeuIysGlnPheLeuIysIyThrIys 139
Db      241 ATCTTCATCATCAGATAGTACGTGTCTATCAGGACCTCCAGCAATTCCTCAAAAAGACAAAG 300
QY      140 LysAsnHisIlyValIleMetAsnAlaArg 148
Db      301 AAGAACCAACAGAGCCATGACGCCCGGTATGAGGAGAGCGGCTGAGGACCAACGAGGAGAC 360
QY      148 148
Db      361 AGAGAGGGGTGGGAGAGCTCGGGGACTCGGAGTGTGAGGGGGTGTGCCCGGCGCTCGG 420
QY      148 148
Db      421 ACAGAGGCTGGGCGAGATGCGGGGCGGCTCCGAGGCCACCGCCTCTCTCGGCCCC 480
QY      149 149
Db      481 ACCGACCGAGAGTCTGCTCGCTCCGCGCCAGAGGCTGAGAGCGCTGGTGACGAGATG 540
QY      158 LeuSerAlaLeuSerPheLeuHisAlaCysSerProPheIleHisGlyAsnLeuThr 177
Db      541 CTGTCTGGGCTCAGCTTCTCTGACAGCCCTGACAGCCCCCAATCATCAGAGGAACCTGACC 600
QY      178 SerAspThrIlePheIleGlnHisAsnGlyLeuIleIysIleGlySerValTrpHisArg 197
Db      601 AGCGACACATCTTATTCAGACAAACGGCTCATCAAGATGAGGTGCGTGAGGACCGA 660
QY      198 IlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIle 217
Db      661 ATCTTCTCAAT-----GACTCTCAAGATGATCTCCGAAGCCCATC 702
QY      218 ArgAlaGluArgGluGlnLeuArgAsnLeuHisPhePheProProGlyIyGluVal 237
Db      703 CGCGCTGACGAGAGAACTTCGAACTTCGACTTCTTCCCCCAAGATGAGAGAGGTG 762
QY      238 AlaAspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaVal 257
Db      763 GCCGATGGAGCCGCTGTGACATCTTCTTTGGAGATGTGTGGCTGGAATGTGCTGTGA 822
QY      258 LeuGluIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArg 277
Db      823 CTGGAATTCAGAACCAATGAGGACACCGGGGTGACAGAGAGGCCATTGCTCGCCACAG 882
QY      278 HisSerLeuSerAspProAsnMetArgGluPheIleLeuCysValLeuAlaArgAspPro 297
Db      883 CACTCGTAGTAGACCCCAACATGCGGAGATTCTTGTGCTGTGCTGCGCGGAGACCT 942

```



```
Db 301 AAGAACCAAGAGCCATTGAACGCCGGGATGGGGAGCGGGGCTGGGGCAAGCCAGGGGAC 360
Qy 148 ----- 148
Db 361 AGACGGGGTTGGGCGAGCCCTCGGGAGTGGATGTGAGGGGTCGCCGGGCTCGG 420
Qy 148 ----- 148
Db 421 ACAAGGGCTGGGCGAGAGATCGGGGGGCGGCTCCGAGGCCAGCCCTCTCTGGGCCC 480
Qy 149 ----- 149
Db 481 ACCCGACGAGCGAGTGTGGTCCGCCGCCAGGCTCGAAGCGCTGGTGCAGCGAGATC 540
Qy 149 ----- 149
Db 541 CTGTCTGCCTCAGCTTCTCTGCAGCGCTGCAGCCCAATCATCCAGGAACTGAC 600
Qy 178 SerAapThrlPheIleGlnHlaAnglYleuIleYlSerProIleIleHlaAnglYlPheIle 197
Db 601 ACCGACACCATCTTCAATTCAGCAACAGGCTTCATCAAGATCGGCTCGTGGCAACGA 660
Qy 198 IlePheSerAsnAlaLeuArgProThrlAlaLeuProAspAapLeuArgSerProIle 217
Db 661 ATCTTCTCCAAAT-----GCACTTCCAGATGATCTCCGAAAGCCCATC 702
Qy 218 ArgAlaGluArgGluGluLeuArgAenLeuHlaPhePheProProGluIlyGluVal 237
Db 703 CCGGCTGAGCGAGAGAACTTCGGAACCTGCACTTCTCCCGACAGATATGAGAGGTG 762
Qy 238 AlaAapGlyThrlValAlaPrlIlePheSerPheGlyMerCysAlaLeuGluMetAlaVal 257
Db 763 GCCGATGGGACCGCTGTGACATCTTCTCTTGGGATGTGGCGGTGAGATGGCTGTA 822
Qy 258 LeuGluIleGlnThrlAsnGlyAspThrlArgValThlGluAlaIleAlaArgAlaArg 277
Db 823 CTGGAATTCACACCAATGGGAGACCGGGTCAAGAGAGAGCCATCTCTCGCCGACAG 882
Qy 278 HisSerLeuSerAapProAenMerArgIupheIleLeuCysAlaLeuAlaArgAapPro 297
Db 883 CACTCGTGAAGTGAACCCCAACATGGCGGAGTTTCACTTGTGCTGGCCGGGACCT 942
Qy 298 AlaAaGArgProSerAlaHisSerLeuLeuPheHlaArgValLeuPheGluValHisSer 317
Db 943 GCCCCCGGCGCTCTGCGCCACAGCCTCTCTTCCACCGGCTCTTCCAGGTGACCTG 1002
Qy 318 LeuLysLeuLeuAlaAlaHisCysPheIleGlnHlaAnglYlSerProIleValAlaVal 337
Db 1003 CTGAAGCTCTTGGCAGCCCACTGCTTCACTGCAGCACCAATACCTCATGCTGAGAAATG 1062
Qy 338 ValGluGluYlThrlYsAlaMerAapLeuHlaValLeuAlaGluLeuProArgPro 357
Db 1063 GTGAGAGAGAAACCAAGGCCATGAGCTCGACGGGTCTTGGCGGAGCTTCCCGGCCC 1122
Qy 358 ArgAaGArgProLeuGlnThrlArgYlSerGluValSerPheMerGluLeuAapLysPhe 377
Db 1123 CGCAGGCGCCCGCTCAGTGGGGACTCGAAGAGCTCTTATGAGAGCTGACAAATTC 1182
Qy 378 LeuGluAapValArgAsnGlyIleYlPrlProLeuMetAapPheAlaAlaThrlArgProLeu 397
Db 1183 CTGAGAGATGTGAGAAATGGAATCTAACCACTGAATGAACTTTCAGAGCCCTGAGACCCCTG 1242
Qy 398 GlyLeuProArgValLeuAlaProProGluGluValGlnLysAlaAlaYlThrlProThrl 417
Db 1243 GGGGCTGCCCGGTGTCTGGCCCAACCCCGAGGAGGTTCAAAGGCCAAGACCCGAGAG 1302
Qy 418 ProGluProPheAapSerGluThrlArgLysValIleGlnMerGlnCysAsnLeuGluArg 437
Db 1303 CCAAGGCCCTTGTGACTGAGACCAAGAAAGGTCACTCAAGATCCAGTGCACCTGAGAGGA 1362
Qy 438 SerGluAapLysValArgTrpHisLeuThrlLeuLeuValLeuGluAapArgLeuHis 457
```

```
Db 1363 AGGAGGACAAAGCGCGCTGGCATCTCACTCTTCTGTGTGCTGAGAAAGCCGCTGCAC 1422
Qy 458 ArgGlnLeuThrlYlThrlAapLeuProThrlAapSerAlaGlnAapLeuAlaSerGluLeu 477
Db 1423 CGGAGACTGACCTACGACCTGCTCCCAAGGAGAGGCCCAAGACTTCGCTCGAGGCTC 1482
Qy 478 ValHisTrpGlyPheLeuHlaGluAapAapArgMetLysLeuAlaAlaPheLeuGluSer 497
Db 1483 GTGCATATGAGCTTCTCTCCACGAGACGACCGGATGAACTGGCCGCTTCTCGAGAGC 1542
Qy 498 ThrPheLeuLysTrpArgIlyThrlAla 507
Db 1543 ACCTTCTCAATGATCCGTGGAGACCAAGGCC 1572

RESULT 12
AAK51824
ID AAK51824 standard; cDNA, 1501 BP.
XX
AC AAK51824;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 369.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
05-FEB-2001; 2001WO-US004098.
XX
03-FEB-2000; 2000US-00496914.
XX
PR 27-APR-2000; 2000US-00560875.
XX
PR 20-JUN-2000; 2000US-00598075.
XX
PR 19-JUL-2000; 2000US-00620325.
XX
PR 01-SEP-2000; 2000US-00654936.
XX
PR 15-SEP-2000; 2000US-00663561.
XX
PR 20-OCT-2000; 2000US-00693325.
XX
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Gao Y;
PI Ma Y, Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
DR MPI: 2001-476283/51.
XX
DR P-PSDB; AAM78691.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 1; Page 1404-1406; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
```

XX Sequence 1501 BP; 300 A; 504 C; 432 G; 264 T; 0 U; 1 Other;
SQ

Alignment Scores:

Pred. No.:	2.97e-212	Length:	1501
------------	-----------	---------	------

Score:	2201.00	Matches:	421
--------	---------	----------	-----

Percent Similarity:	98.37%	Conservative:	2
Percent Identity:	97.01%	WGA:	1

Best Local Similarity:	97.91%	Mismatches:	1
Query Match:	82.43%	Indels:	5

Query match:	82.43%	Index:	6
DB:	4	Gaps:	1

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

US-10-618-941-99 (1-507) x AAK51824 (1-1501)

OY	78	LYSVALPHEVALAIAHISGLIULYVILEGINTHVALPHEGLIULNEVALLEUVAL	97
Db	55	CGTGCCTTCGGGCGCACAGAGGAAGATTCACACCGTGTTCAGAGAGCGTGGTGGTG	114
OY	98	APRHISPROAMILEVALYSLYVITYRTPLLEUAPRHISERGIUALIACYSAIA	117
Db	115	GACCAACCCGACATCTGTGAAGTTGCACAAAGTACTGGCTGGATTCCTCTAGAGCGCTGGC	174
OY	118	ARGVALLEPHEILETHRGUITYVALSERSEGLYSERLEULYGLNPHLEULYLYS	137
Db	175	AGGGTCATCTTCATCCACAGATACGTGTATCATCGGAGCGCTCAAGCAATTCCTCAAAAAG	234
OY	138	THIRLYLYSASNIILYVALIAMELANAIAARGVALITRPLYSAIRGTPYSTRHINILE	157
Db	235	ACCAAGAGAACACAAAGCGCATGAAGCGCCGGGCGCTGGAAAGCGTGGTGACAGCAATC	294
OY	158	LEUSERRIALAUSERPHELEULIYALACYSERPROPROILEILEHISGLIYALEULTHR	177
Db	295	CGTGTGCCTCAGCTTCCTGCACCGCTCAGCCCCCATCATCCACGGGAACCTGACC	354
OY	178	SERAPRTHILPHEILEGNIISAENGILYULIILEYALIEGLYSERVALITPHISARG	197
Db	335	ACGGACACCATCTTCATTCAGACAAAGGCTCATCAAGTACGGCTCCGTTGGACCGA	414
OY	198	ILEPHESEASNALALEUARGPROPROTHRALALEUPROASPARPHEUARGSERPROILE	217
Db	415	ATCTTCTCCAAAT-----GCATCTCCAGATGATCTCCGAAGCCCCCATC	456
OY	218	ARGVALAGUARGGLULYUARGANILEHISPHERPHEROPROGLIUTYRGILYGLUVAL	237
Db	457	CGCGCTGACCGAGAGAACTTCGGAACCTGCACATTCTTCCCCCAAGATATGGAGAGGTG	516
OY	238	ALAASRGLYTHRALVALASRIIPHESERPHEGLYMERCYVALALEUGLUMETALAVAL	257
Db	517	GCGGATGGGACCGGTGTGACATCTTCTCTTTGGGATGTGTGCTGTGAGAGTGGCTGTGA	576
OY	258	LEUGLUILEGLINTHRAENGILYAPRTHIRATGVALTHRGUIGUALIILEALAAARGAARG	277
Db	577	CTGGAAATCCAGACCAATGGGGACACCGGGGTCCACAGAGAGGCCAATGTCTCCGGCCAGG	636
OY	278	HISSERLSESRASRPROANMECTARGLUPHEIILEUCYSECYSEULIAARGAAPRO	297
Db	637	CACCTGGCTAGTGAACCCCAACCATGGCGGAGTTATCTTCTGCTGCTGGCCCGGAGACCT	696
OY	298	ALAARGARGPROSERALAHISSERIEULEUPHENIARGVALLEUPHEGLIUALHISER	317
Db	697	GCCCCCGGCGCTCTCCACAGCGCTCTTCCACCGCGGTCTTGGCGGAGTCTCCCGGCGCC	756
OY	318	LEULYVLEULEUALAHISCYSPHEILEGNIHISGLINTYRTLEUMETPROGLIUMANVAL	337
Db	757	CTGAAGCTCTTCGACGCCCACTGTCTTCATCCACACCAAGTACTCAATGCTTGAAATGTG	816
OY	338	VALIGLULIYESTHRILYVALAMEASRPLEULHISALIVALLEUALIGLULEUPROARGPRO	357
Db	817	GTGGAGAGAAAGACAAAGCGCATGACCTGACAGCGGTCTTGGCGGAGTCTCCCGGCGCC	876
OY	358	ARGARGPROPROLEUGINTPRATGYRSEGLIUALISERPHENEGULYUABRILYPHE	377
Db	877	CGCAGGCCCCCGCTGACGTGGCGGTACTCGGAAGTCTCTTCATGGAAGCTGGCAAAATTC	936

QY	378	LeuGIuAaPvAlArYAsnGIyIeYrProLeuMeLsnPheAlAlaThraArgProLeu	357
Db	937	CTGGAGGATGTCAGAAATGAATCTCAACCACTGAATGAACTTGGACCCACTCAACCCCTG	996
QY	398	GIYLeuPArYArgValLeuAlaProProProGIuGIuValGIuLbYalYalYrThrProThr	417
Db	997	GAGCTGCCCCGTGTCTGTGCCCCCAACCCCGAGAGAGTCCAAAGGCCAAAGACCCGAGC	1055
QY	418	ProGIuProPheAspSerGIuThraArgYValIleGIuMetGIuCYsAsnLeuGIuArg	437
Db	1057	CCAGAGCCCTTTGACTGTAGACCAAGAAAGTCATCAATCAATCAAGTCAACTGGAGAGA	1111
QY	438	SerGIuAaPvAlArYArgTrpHisLeuThrLeuLeuValLeuGIuAaPArgLeuHis	457
Db	1117	AGCGAGGACAAAGCCCGCTGGCATCTCACTGCTTGTGTCTGTGAACACCGCTGCAC	1177
QY	458	ArgGIuAlaThrThyAspLeuLeuProThrAspSerAlaGIuAspLeuAlaSerGIuLeu	477
Db	1177	CGGCAAGTGAACCTTAGCACTGTCTCCCAAGGCAAGGCCCAAGACTCGCTGTGGAGCTC	1233
QY	478	ValHisTrYGlyPheLeuHisGIuAaPAspArgMetYLeuAlaAlaPheLeuGIuSer	497
Db	1237	GGGCACTATGGCTTCTCTCCACGACGACGACCGGATGAACCTGGCGCTTCTCGAGAGC	1293
QY	498	ThrPheLeuYrTYrArgGIYThrAla	507
Db	1297	ACCTTCTCAAGTACCGTGGACCCAGAGGC	1326

RESULT 13
 AAF29899
 ID AAF29899 standard; DNA, 981 BP.
 AC AAF29899;
 XX
 DT 04-APR-2001 (first entry)
 XX
 DE Human h15993 DNA.
 XX
 KW Human; protein kinase; cell growth; tumour; cancer; immune; inflammatory;
 KW respiratory; haematological; bone disorder; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200100879-A1.
 XX
 PD 04-JAN-2001.
 XX
 PF 30-JUN-2000; 2000WO-US018291.
 XX
 PR 30-JUN-1999; 99US-00345473.
 PR 01-MAY-2000; 2000US-00562480.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Hodge MR, Meyers R, Williamson M;
 DR WPI; 2001-061977/07.
 XX
 PT New protein kinase polypeptides, nucleic acids and anti-kinase
 PT antibodies, useful for diagnosing and treating e.g. cancer, inflammatory,
 PT immune, cardiovascular and bone disorders.
 XX
 PS Claim 1; Fig 13; 93pp; English.
 XX
 CC The present invention relates to human protein kinase. The proteins are
 CC from new human genes termed h12832, h1418, h14633, h15990, h15993,
 CC h16541 and h2552. The proteins may be used to identify modulators of
 CC their activity. The proteins may also be used to derive products for the
 CC treatment of cellular growth related disorders, malignancies, cancer,
 CC immune, inflammatory, respiratory, haematological and bone-related
 CC disorders
 XX
 SQ Sequence 981 BP; 209 A; 298 C; 309 G; 162 T; 0 U; 3 Other;

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 13, 2006, 21:11:56 ; Search time 4609 Seconds
(without alignments)
5146.681 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670
Sequence: 1 MAAPAPAPRARRERERED.....RMKLAAPLESTFLKRYGTQA 507

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1p
-Q/cgnt2_1/USPTO.epool_p/US10618941/runatc_12012006_070204_14281/app_query.fasta_1.647
-DB=EST -QPMT=fastcap -SUFFIX=ext -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdd
-DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMF=p2n -NORW=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10618941.CCGN_1.1.8010.0/runatc_12012006_070204_14281 -NCPU=6 -ICPU=3
-NO MMMP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WANN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_g881:*
10: gb_g882:*
11: gb_g883:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2391	89.6	3163	4 BC068117	Mus muscu
2	2037.5	76.3	3729	4 BC037396	Homo sapi
3	1562.5	58.5	2163	4 AK046142	Mus muscu
4	1537.5	57.6	1608	10 AY408905	Mus muscu
5	1536	57.5	2616	4 HSM605653	AL834530 Homo sapi
6	1534.5	57.5	1601	10 AY408904	Pan trogl
7	1534.5	57.5	1608	10 AY408903	Homo sapi

8	1534.5	57.5	2137	4 CR604477	full-leng
9	1534.5	57.5	2158	4 HSM601650	Homo sapi
10	1534.5	57.5	2181	4 CR858659	Pongo pyg
11	1520	56.9	2086	4 CR591118	full-leng
12	1518.5	56.9	2181	4 AK077440	Mus muscu
13	1481	55.5	899	7 CR998804	CR998804
14	1403	52.5	927	5 BX410189	BX410189
15	1429	46.0	825	3 BI645887	BI645887
16	1199	44.9	729	6 CD354461	CD354461
17	1170	43.8	2923	4 AF318376	AF318376
18	1165.5	43.7	859	7 CK476491	CK476491
19	1161	43.5	673	8 DN998004	DN998004
20	1160.5	43.5	841	7 CO248961	CO248961
21	1126	42.2	716	7 CV105566	CV105566
22	1109	41.5	700	7 CN358568	CN358568
23	1100.5	41.2	1028	1 AL532940	AL532940
24	1098	41.1	899	2 BF299723	BF299723
25	1055.5	39.5	770	7 CN218030	CN218030
26	1050.5	39.3	725	2 BG913260	BG913260
27	1042	39.0	651	6 CB557694	CB557694
28	1030	38.6	996	3 BI650160	BI650160
29	1025.5	38.4	911	3 BU508281	BU508281
30	1015	38.0	959	5 BQ652622	BQ652622
31	1013	37.9	643	7 CN358567	CN358567
32	1012.5	37.9	1382	8 DN656596	DN656596
33	1011	37.9	898	5 BQ649309	BQ649309
34	996	37.3	897	5 BQ839694	BQ839694
35	989	37.0	869	8 CX566377	CX566377
36	988.5	37.0	929	8 CX917126	CX917126
37	986	36.9	1001	5 BQ716004	BQ716004
38	978	36.6	637	7 CN790575	CN790575
39	971.5	36.4	922	7 CR982195	CR982195
40	969.5	36.3	1122	3 BI410402	BI410402
41	968.5	36.3	756	2 BF134441	BF134441
42	966	36.2	730	6 CB326067	CB326067
43	964.5	36.1	950	5 BQ648254	BQ648254
44	960	36.0	586	6 CD354720	CD354720
45	956	35.8	819	8 CX907573	CX907573

ALIGNMENTS

RESULT 1	BC068117	3163 bp	mRNA	linear	HTC 08-FEB-2005
LOCUS	BC068117				
DEFINITION	Mus musculus nuclear receptor binding protein 2, mRNA (cDNA clone IMAGE:6419290).				
ACCESSION	BC068117				
VERSION	BC068117.1	GI:45829694			
KEYWORDS	HTC.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1 (bases 1 to 3163)				
AUTHORS	Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heide,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stاپleton,M., Soares,M.B., Donald,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Utsid,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Morley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettlemen,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G., Blakeley,R.W., Touchman,D.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalske,U., Smallus,D.E.,				

TITLE Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL 12477932
 PUBMED 2 (bases 1 to 3163)
 REFERENCE Director MGC Project.
 AUTHORS Direct Submission
 TITLE Submitted (25-MAR-2004) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgabbs-rc@mail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.
 Thomas L. Casavant.
 Web site: <http://genome.uiowa.edu>
 Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu
 Bonaldo, M.F., Akabogu, I., Bair, T., Bair, J., Crouch, K., Davis, A.,
 Fisher, K., Keppel, C., Kucaba, T., Lebeck, M., Melo, A., Schaefer, K.,
 Scheetz, T., Smith, C., Snir, E., Tack, D., Trout, K., Walters, J.,
 Casavant, T., Soares, M.B.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: Plate: Row: Column: 0
 This clone has the following problem: clone inconsistent with known
 gene structure

FEATURES
 source
 Location/Qualifiers
 1..3163
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:6419290"
 /issue_type="Brain, mouse 15.5 dpc"
 /clone_lib="NIH BMAP_EW0"
 /lab_host="DH10B"
 /note="Vector: pYX-ASC"

ORIGIN
 Alignment Scores:
 Pred. No.: 5,74e-243 Length: 3163
 Score: 2391.00 Matches: 470
 Percent Similarity: 85.44% Conservative: 11
 Best Local Similarity: 83.48% Mismatches: 13
 Query Match: 89.55% Indels: 70
 DB: 4 Gaps: 4

US-10-618-941-99 (1-507) x BC068117 (1-3163)

Db 247 GGGGCCCATGAGAGAGATCCAGCATGTTTGAAGACGTGGCGGTGAGCATCC 306
 QY 101 Aaalllvallyleuuhlslystyttrpleuapthrserrlualcyalaargvallle 120
 Db 307 AACATCGTCAGACGTGACAAAGATTGGCTGAGCGCTTGAAGGCCGCGCAG- 359
 QY 121 Pheillethrglutyvalserserglsyserleuylvglnphleuulysythrlyslvs 140
 Db 360 -----GAATAGGTGTGTCTGGACCTTCAGACAGTCTTCAAAAAGCAAGAAAG 410
 QY 141 AaahlslyvalameAaahlaarg----- 148
 Db 411 AACCAAGAGCATGAGCGCCGGATATGGGAGTGGCGTGGGACGTGAGAGATCCAG 470
 QY 148 ----- 148
 Db 471 CAGAACTTAGGACAGTAGTGGGGGAGACTCGAAAAACCTCGATTCAAGATGGGG 530
 QY 148 ----- 148
 Db 531 TAGCGGGCGGGCTCCGACAGCTCAGCAGCTTTGGTCCACCCGACGAGTGTGGCT 590
 QY 149 -----AATrPLysArgTrpCythrGlnlleuSerAlaleuSerPheleuh1 165
 Db 591 CGGCTGCCAGGCTCGGAAGCGCTGGGTATCGAGATCCGTCTGCACCTAGTTTTCGA 650
 QY 165 AAlAcYserPProProllelleHleGlyAauleuThrsAappThrllePhe1leGlnH1 185
 Db 651 CCCCTGCAGTCCCCCATCATCACCAGGAACCTGACGACGACACCATCTTCATTACAGA 710
 QY 185 sAaHnGlyleuilelyslleGlyserValThrlsAargllePheSerAaahlaAleuAargr 205
 Db 711 CAATGGCTCATCAAGATCGCTCCGTGTGTAACGACATCTTCCAAAT----- 759
 QY 205 oProThralaleuProAappPleuAargSerProilleaarglaAgluarGluuleuar 225
 Db 760 -----GCACCTTCTGATGATCTCCGAGCCCTATTCAGAGCTGAACGGGAAGACTCCG 812
 QY 225 gAaHleuuhlsPhePheProProGluTrgIyGluValAlaAapGlyThrlAlaValAap11 245
 Db 813 AAACCTGCACCTTTTCCACACAGAGTATGGCAAGTCAATGATGAGGACTGTGTGACAT 872
 QY 245 ePheSerPheGlyMetCyalaAleuGluMeKAlaValleuGluuileGlnThrsAaHlyvs 265
 Db 873 CTTCCTTCCTGGAGATGTGCACTGAGATGCTGATCTGAGATCCAGATCCAGCCAG 932
 QY 265 pThrlArgValThrgluGluAla1leAlaAarglaAargHlsSerleuSerAappProAaHke 285
 Db 933 TACCAAGTCAAGAAAGGCCATCGCTGACAGCCACTGATGATGAGACCCCAACT 992
 QY 285 rArgGluPheilleuCyAseleuAlaAargAappProAlaAargAargProSerAlaHsse 305
 Db 993 GCGGAATTCATCTCTCTGCTGCGCGGAGCCCTGCGCCGACCTCCAGCCACACAA 1052
 QY 305 rleuSeuPheHlsAargValleuPheGluValHlsSerleuylseuAaHlsAaHlycy 325
 Db 1053 CCTCTCTTCCACCAAGATGCTCTTGAAGTGACCTGCGAAGCTGTGCGACGCTACATG 1112
 QY 325 sPheilleGlnHlsGlnTryleuMeCProGluAaahVal1AgluGluylsThrlsAaHme 345
 Db 1113 CTTCAATCCAGCACCAAGTACTCATGCTGAGATGTGTAAGAGAAAGAAACCAAGGCCAT 1172
 QY 345 rAapleuuhlsAlaValleuAlagluSeuProAargProAargAargProProleuGlnTrpr 365
 Db 1172 GGAACCTCCATGAGTTTGGCTGAGATGCGCAGCCCATGAGACCCCAATCAAGTGGCG 1232
 QY 365 gTyrSerGluValSerPheMeGluLeuAaplyPheleuGluAapValArgAaHly11 385
 Db 1233 GTACTCAGAGGTCTCTCTTGGAGCTGACAAATTCCTTAAGAGATGTTCAGAAACGGAGT 1292
 QY 385 eTyrProleuMeCaapAaHlsAla1ThrsAargProleuGlyleuPProAargValleuA1ar 405
 Db 1293 CTAATCAGATGAACCTTGGCGGTGCTGCGCCCTTGGGGCTTCCCGTGTGGCCCC 1352

QY	405	oPrcProBiolGuValaGlnuValaAlaVThProthProGluuProPheApsSerGluTh	425		
Db	1333	ACCCCAAGGAAAGGCCCAAAAGGCCAAACCTCAACGGCAAGACCTTTGACTCGGAAC	1412		
QY	425	rArgIyValIleGlnMeGlnCyAsnIleuGluuArgSerGluuAspIyValaArgTrpHi	445		
Db	1413	CAGGAAGGTGGTCCAGATGCGATGCGACCTGGAAAGAGCGAGGACAAAGGCTCGGTGCA	1472		
QY	445	bleuThIreuleuIeuValIleuGluuAspArgIeuHiArgGlnIeuThrTyraAspIeu	465		
Db	1473	CCTTACTCTGCTCTTGGCTTGGACGACCGGCTTACATCGGACGCTGACTTGAATCTGCT	1532		
QY	465	uProthrApsSerIaGlnAAspIleuAlaSerGluIeuValuHisTrArgIyPheIeuHiArgI	485		
Db	1533	CCCAACGGAACAGTGGCCCAAGACCTTGGCTGTGAATAGCATTTATGAGCTTTCCTGCA	1592		
QY	485	uAspAspArgMeIyIeuAlaAlaAphaIeuGluSerThrPheIeuIyTyraArgIyTh	505		
Db	1593	GGATGACAGGACCAAGGCTTAGCGCTTCTTGAGACCACTTTTCTCAAGTACCGAGGAC	1652		
QY	505	rgIuAla	507		
Db	1653	GCAAGCG	1659		
RESULT 2	BC037396	3729 bp	mRNA	linear	HTC 05-APR-2005
LOCUS	BC037396	3729 bp	mRNA	linear	HTC 05-APR-2005
DEFINITION	Homo sapiens nuclear receptor binding protein 2, mRNA (cDNA clone IMAGE:5180619), with apparent retained intron.				
ACCESSION	BC037396				
VERSION	BC037396.1	GI:23337035			
KEYWORDS	HTC.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 3729)				
	Strauberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buelow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Datschenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stepien, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulhaly, S.J., Bosak, S.A., McKwan, P.J., McEwan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Wille, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Viallon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bonfield, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Gilmwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywicki, M.I., Skalka, U., Smalins, D.B., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
CONSTRM	Mammalian Gene Collection Program Team				
TITLE	Human and mouse CDNA sequences of more than 15,000 full-length				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 3729)				
AUTHORS	NIH MGC Project				
CONSTRM	Direct Submission				
TITLE	Submitted (04-SEP-2002) National Institutes of Health, Mammalian				
JOURNAL	Gene Collection (MGC), Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov				
COMMENT	Contact: MGC help desk Email: cgmpbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc.				

		cDNA Library: Arrayed by: The I.M.A.G.E. Consortium (LILN)	
		DNA Sequencing by: Baylor College of Medicine Human Genome	
		Sequencing Center	
		Center code: BCM-HGSC	
		Web site: http://www.hgsc.bcm.tmc.edu/cdna/	
		Contact: amgobcm.tmc.edu	
		Gunnarane, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulaged, H.,	
		Kovis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,	
		A.N., Gibbs, R.A.	
		<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LILN at: http://image.llnl.gov</p> <p>Series: IRAK Plate: 50 Row: b Column: 20</p> <p>This clone has the following problem: retained intron.</p>	
FEATURES	source	<p>Location/Qualifiers</p> <p>1..3729</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:5180619"</p> <p>/tissue_type="Brain, Lung, Testis, adult, pooled whole"</p> <p>/clone_id="NH MGC_115"</p> <p>/lab_host="DH10B"</p> <p>/note="Vector: pCMV-SPORT6"</p>	
ORIGIN			
Alignment Scores:			
Pred. No.:	3,74e-205	Length:	3729
Score:	2037.50	Matches:	415
Percent Similarity:	75.59%	Conservative:	0
Best Local Similarity:	75.59%	Mismatches:	1
Query Match:	76.31%	Indels:	134
DB:	4	Gaps:	4
US-10-618-941-99 (1-507) x BC037396 (1-3729)			
QY	86	LYSLIEGINTTrrvalpHeGlUGlnleuValaaphisProAsnllleValysleu	105
Db	3	AAGATCCAGACCGGTTCGAGCAGCTGTGTGGTGAGCACCCGAAATCGTGAATTG	62
QY	106	HlslysrTYrrTPLeuAspThrSerGlnlaCysAlaTg-----	118
Db	63	CACAAgTACTGGCTGGATACCTCTGAGGCTTGGCGAGGGTGACACGGGCAAGCCCCG	122
QY	118	-----	118
Db	123	GGCAGGCAAGAACCACTAGTGGGGCGGGCATGGGAGCAACCCGACGCTGTGAA	182
QY	119	-----VallelPheIleThrgIuTyrrValSerSerGlySerle	131
Db	183	ATTCCTGAGCTGCACCGTGCAGAGTCATCTTCATCACAAGATACGTGCATCAGCAGCCT	242
QY	131	uLyGlnPheleuylsYrThrrlyblyAsnHlilyAlaMetAlaAlaTg-----	148
Db	243	CAAGCAATCTCTCAAAAAGACCAAGAAACCAAGGACCAATGAAAGCCCGGGTATGGGG	302
QY	148	-----	148
Db	303	AGCGGGCTGGGGCAGCCACGGGAGACGAGACGGGGTGGGGCAGCTCGGGACATGGGATG	362
QY	148	-----	148
Db	363	GTGAGGGGGTGCCTCGGCGGCTCGGACAGGGGCTGGCGAGATGCGGGCGGGCTCGCG	422
QY	149	-----Al	149
Db	423	AGGCCGACCGCTCTCTCGGCGCCACCCGACCGAGTCTGTCGCGCGGCGAGGC	482
QY	149	ATrPlsArTgTrpCysThrglnlleleuSerAlaIleuSerPheleuHlsAlaCysSerPr	169
Db	483	CTGGAAGGCTCGTGACGACGATCTCTGTCTGCGCTTCTCTGCACGCTCGAGACCC	542
QY	169	optolIleHlsGlyAsnleuThrSerAspThrIlePheIleGlnHlsAmGlyleuHl	189

Db 543 CCAATCATCTCAGGGAACCTGACAGCAGCACCATTTTATTCAGACACAGCCCTCAT
 Qy 189 elyyllelylserValTrpHlsargllepheserAenAlaLeuArgProThrAlaIe 209
 Db 603 CAAGATCGGCTCGGTGGCAGCGAATCTTCCAAAT-----CCACT 644
 Qy 209 uProAapAapLeuArgSerProIleArgAlaGluuArgGluLeuArgenleuHlspr 229
 Db 645 TCCAGATGATCTCGAAGCCCCCATCCGCTGAGCGAGAGAACTTCGAACTCGACATT 704
 Qy 229 ephProProGluuTyrglyGluValAlaAaprglyThAlaValAapIlepheserPhegl 249
 Db 705 CTTCCCCCAAGATGATGAGAGGAGGCGCATGGGACCGCTGTGACATCTTCTTGG 764
 Qy 249 yMeCyAlaLeuGluMeAlaValleuGluIleGlnThrAenGlyAapThArgValTh 269
 Db 765 GATGTGTGCTCGAGAGTGGCTGTACTGAAATCCAGAACAAATGGGACACCCGGGTAC 824
 Qy 269 rglugluAlaIleAlaArgAlaArgHlsSerleuSerAapProAapMetArgIleHls 289
 Db 825 AGAGAGAGCCATGCTCGCCAGGCACTGCTGAGTGAACCCAACTGCGGAGTTCAAT 884
 Qy 289 eleuCyCyAlaLeuAlaArgAapProAlaArgAapProSerAlaHlsSerleuAenHls 309
 Db 885 CTTTGTGCTGCTGGCCCGGAGCCCTGCCCCCGGCTCTGCCACAGCTCTTCTTCCA 944
 Qy 309 sarValleuPheGluValHlsSerleuTyAlaAlaHlsPheHlsleGlnHls 329
 Db 945 CCCCCTGCTCTTCGAGGTGCATCTGCTGAGACTCTTGAGCCCACTGCTTCATCCAGCA 1004
 Qy 329 sglNTrleuMetProGluAenValValGluGluTyThAlaMetAapleuHlsHls 349
 Db 1005 CCAGTACTATCTGCTGAGATGTGTGGAGGAGAAACCAAGCCATGACCTGACGCG 1064
 Qy 349 aValleuAlaGluLeuProArgProArgAapProProleuGlnTrpArgTySerGluVal 369
 Db 1065 GGTCTTGCGGAGCTTCCCGGCCCCGAGGCCCCCGCTGCAATGGCGGTACTCGAAGT 1124
 Qy 369 lSerPheMetGluLeuAapTyPheleuGluAapValArgAenGlyIleTyPProleuMe 389
 Db 1125 CTCCTTCATGAGCTGGACAAATTCCTGAGAGATGTGAGAAATGGAATCAACCATGAT 1184
 Qy 389 lAanPheAlaAlaThrArgProleuGlyLeuProArgValleuAlaProProProgluG 409
 Db 1185 GAACCTTGCAGCAGCTGACCTGAGGCTGCCCGGTGTGCTGCCCCACCCCGAGAGA 1244
 Qy 409 uValGlnTyAlaIleTyThProThrProGluuProPheAapSerGluThrArgHlsValI 429
 Db 1245 GGTCCAAAGGCCAAGACCCCGAGCGCAAGCCCTTTGACTTGAAGHCCAGAAAGTCAAT 1304
 Qy 429 eglMetGlnCyAaenleuGluArgSerGluAapTyAlaArgTrpHlsleuThrleuLe 449
 Db 1305 CCAGATGCACTGCAACTGGAGAGAGAGCGAGCAAGCGCGCTGGCATCTCATCTCGCT 1364
 Qy 449 uLeuValleuGluAapArgleuHlsArgGlnleuTrpTyAapleuLeuProThrAapS 469
 Db 1365 TCTGTGTCTGGAAGACCGGTGACCGGAGCTGACCTTACGACTCTCCCAACGAGCAG 1424
 Qy 469 rAlaGlnAapLeuAlaSerGluLeuValHlsTyrglyPheleuHlsGlu----- 485
 Db 1425 CGCCCAAGACCTGCGCTCGAGACTGTGTGACTTATGGCTTCTTCACAGA-GGTGCGCTGG 1483
 Qy 485 ----- 485
 Db 1484 CGGTGCGGCGGCGCTGCGAGAGGGCGCAGCGGGGCGCGGCCCTCGTCCCCCA 1543
 Qy 486 -----AapAapArgMetTySleuAlaAlaPheleuGluSerThrp 499
 Db 1544 TGCCTCCTCTCTTCCGAGAGACACCGAGTGAAGTGGCCGCTTCTCGAGAGACACTT 1603
 Qy 499 eleuTyTyArgGlyThrGlnAla 507
 |||||

Db 1604 CCTCAAGTACCGTGGAGCCAGGCC 1628
 RESULT 3
 AK046142
 LOCUS
 DEFINITION
 AK046142 2163 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult male corpora quadrigemina cDNA, RIKEN
 full-length enriched library, clone:B230344117 Product:SIMILAR TO
 NUCLEAR RECEPTOR BINDING PROTEIN (HLR-INTERACTING PROTEIN KINASE)
 homolog [Mus musculus], full insert sequence.
 ACCESSION
 AK046142.1 GI:26337852
 VERSION
 KEYWORDS
 HTC, CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P., and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 PUBMED
 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 PUBMED
 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishue, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsushita, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 TITLE
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 PUBMED
 11076861
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the
 PANTOM Consortium.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409, 685-690 (2001)
 REFERENCE
 AUTHORS
 5
 The PANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 TITLE
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 JOURNAL
 Nature 420, 563-573 (2002)
 REFERENCE
 AUTHORS
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
 Nakamura, Y., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohsato, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sasaki, C., Sakai, K., Sakazume, N.,
 Sano, H., Sawaki, D., Shibata, K., Shingawa, A., Shitaki, T.,
 Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,
 Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
 Muramatsu, M. and Hayashizaki, Y.
 TITLE
 Direct SubMISSION
 Submitted (16-JUL-2001) Yoshinide Hayashizaki, The Institute of
 Physical and Chemical Research (RIKEN), Laboratory for Genome
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
 RIKEN Yokohama Institute, 1-7-22 Suhei-cho, Teurumi-ku, Yokohama,
 Kanagawa 230-0045, Japan (E-mail: genome-res@gscl.riken.jp,
 URL: http://genome.gsc.riken.jp/, Tel:81-45-503-9222,
 Fax:81-45-503-9216)

QY 468 AspSerAlaGlnIleuValSerGluLeuValHisIleTyrGlyPheLeuHisGluAspAsp 487
Db 1572 GAAGACATCCCGGACCTTGGACGCTGGAGCTGGGCACTGGCTTCACTAGTGAAGCTGAT 1631

QY 488 ArgMetIleuValAlaPheLeuGluSerThrPheLeuIleTyr 502
Db 1632 CAAAGCCGCTGAGTCTGTCTGTGAGAGACACCTCAACAAGTTC 1676

RESULT 4
AY408905 1608 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus NRP gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY408905.1 GI:39764873
VERSION GSS.
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1608)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeleo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adam, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene clusters
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
AUTHORS 2 (bases 1 to 1608)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
Todd, M.A., Tanenbaum, D.M., Civeleo, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adam, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..1608
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1608
/gene="NRP"
/locus_tag="HCM3379"

ORIGIN
Alignment Scores:
Pred. No.: 2,17e-152 Length: 1608
Score: 1537.50 Matches: 307
Percent Similarity: 73.15% Conservative: 58
Best Local Similarity: 61.52% Mismatches: 105
Query Match: 57.58% Indels: 29
DB: 10 Gaps: 6
US-10-618-941-99 (1-507) x AY408905 (1-1608)

QY 17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyAsp 36
Db 130 GAGGAAGAAGAAAGAGTGAAGTGAAGTCTTGGAAAGACACCTGTGGGCGC 189

QY 37 TTPGlnIleArgArgGluGlnValaGlnGlnIleAspMetProGluLeuGlnSerThrPhe 56
Db 190 TGGCAAGAGGGGCAAGAGGTGAATCAAGCTTAATGTACAGGCACTGACAGTGCATAC 249

QY 57 LeuAlaMetAspThrGluGluGlyValaGluValaIleTTPAsnGluLeuHisPheGlyAsp 76
Db 250 CTGGCCATGGAATACAGAGGAAGGTGTGGAGTGTGTGGAAATGAGGTACAGTTCTGTAA 309

QY 77 ArgIleAlaPheAlaAlaHisGluGluIleGlnThrValPheGluGlnLeuValLeu 96
Db 310 CGCAAGACTACAAACTGCAGAGAGAAAGGCTCGGAGGTGTTATATATGTAATCAG 369

QY 97 ValAspHisProAsnIleValIleuValIleuHisIleTyrTyrIleuAspThrSerGluAlaCys 116
Db 370 CTGGAACATCTTAACCTTTTAAGTTTCAAAATATATGGCTGATTTAAAGAAACAG 429

QY 117 AlaArgValIlePheIleThrGluIleuValaSerSerGlySerIleuValGlnPheLeuVal 136
Db 430 GCTAGGGATTTTTCATCAGAAATATCATGTCTCTGGAGAGTCTTAAGCAGTTTCTGAAG 489

QY 137 IyeThrIleValaAsnHisIleValaMetAlaAlaArgAlaIleTyrIleValaArgTyrPheThrGln 156
Db 490 AAGACCAAAAGAACCAAGCAAGACTGAATGAAGAGCTTGGAAAGCTGTGTATACAG 549

QY 157 IleLeuSerAlaLeuSerPheLeuHisIleValaCysSerProProIleIleHisGluValaLeu 176
Db 550 ATCTCTCTGCGCTTAAGTACTGCACTCTGTGACCTTCCCATTCATCGAGAACTCG 609

QY 177 ThrSerAspThrIlePheIleGlnHisaGlnIleValaIleValaIleGlySerValaTyrPheHis 196
Db 610 ACCTGTGACACCATCTTTCATCCAGACACAGGACTCATCAAGATTGGCTGTGG----- 663

QY 197 ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerPro 216
Db 664 -----GCTCTGACACTATCAACAATTCAC 687

QY 217 IleArgAlaGluArgGluGluIleuArgAsnLeuHisPhePheProProGluIleuValaGlu 236
Db 688 GTGAAGACTTGGCGGGAGAGACAGAAAGCTTACCTTTTGGACACCAAGATGAGAA 747

QY 237 ValAlaAsp---GlyThrAlaValaAspIlePheSerPheGlyMetCysAlaLeuGlnMet 255
Db 748 GTCAACAAGCGTAGAACACAGACAGTGCATCTCTTGGCATGTGTGCATCGAGAGATG 807

QY 256 AlaValaIleGluIleGlnThrArgAsnIleuValaAspThrArg---ValThrGluGluAlaIleAla 274
Db 808 GCAGTGCTGAATTCAGGACAGATGCGAGTCTCATATGTGCGACAGAGAACCATCAC 867

QY 275 ArgAlaArgHisSerIleuSerAspProAsnMetArgGluPheIleuValaCysValaAla 294
Db 868 AGTGCATCCACTACTAGAACATCATTAACAGAGGAGTTTATCAAAAGTCCCTGCAG 927

QY 295 ArgAspProAlaArgArgProSerAlaHisSerIleuPheHisArgValaLeuPheGlu 314
Db 928 TCTGAGCCTGTCTGGAGACCAACAGCAGAGAACTTGTGTCCACCCAGACACTGTGA 987

QY 315 ValHisSerIleuValaLeuValaAlaHisCysPheIleGlnHisGlnIleTyrLeuMetPro 334
Db 988 GTGCCCTCACTCAAGCTTCTGTCTGTCACTGTATCTGTGGGACCCAAACATGATCCCA 1047

QY 335 GluAsnValaIleGluIleuValaMetAspLeuHisAlaValaLeuAlaGluIle 354
Db 1048 GAGAAAGCCTTGAAGAGATATCAAGAAACATGTGATCCAGAGTCTATTAAGTAAAT 1107

QY 355 ProArg---ProArgArgProProLeuGlnIleTyrArgTyrSerGluValaSerPheMetGlu 373
Db 1108 CCCGACAGGCGCAGAGAACAGATTCAGACTTGTACTCTCAGTCACACAGCCCTAAGA 1167

QY 374 LeuAspIleValPheLeuGluAspValaIleArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 1168 TTAGACAAATTCCTTGAAGATGTCAAGATGGAATCTTACCCCTGTGACAGCTTT----- 1221

QY 394 ThrArgProLeuGluIleuProArgValaLeuAlaProProProGluGluVala----- 410
Db 1222 -----GGGCTACTTGGGCTCAGCAGCCACACAGACAGAGAGGTGATCATCCT 1269

QY 411 -----GlnIleAlaIleValaThrProThrProGluProPheAspSerGluThrArgIleVal 427
Db 1270 GTTGTGCCCCCTCTGTCAAGATCTCAACTCTTGAAGCAGCTGAAGATGAGACAGAAAG 1329

QY 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspIleAlaArgTyrPheIleuThr 447

Db 1330 GTGTGCTGATGACAGTGCATCATCTGTGAGGAGGAGGATCAACACCATCTAACA 1389
Qy 448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrArgPheLeuProThr 467
Db 1390 CTTTGTCTGAGAGCTGGAGGACAAATTGAACCGGACCTGACCTGTGATCTGATCCAAAT 1449
Qy 468 AspSerAlaGlnAspLeuAlaSerGluLeuValHisTyrGlyPheLeuHisGluAspAsp 487
Db 1450 GAGAGCATCCCGGACCTGGCAGCTGAGCTGATGACGCTGGCTTCATTAGTAGAGCTGAT 1509
Qy 488 ArgTyrLeuLeuAlaAlaPheLeuGluSerThrPheLeuTyrTyrArgGlyThrGln 506
Db 1510 CAGAGCGCGCTGACTCTCTGCTGAGGAGACGCTCAACAGTTCACTTCAACACGAG 1566
RESULT 5
HSM805653 2616 bp mRNA linear HTC 22-SEP-2004
LOCUS Homo sapiens mRNA, CDNA DKFZp434I2411 (from clone DKFZp434I2411).
DEFINITION AL834530
ACCESSION AL834530.1 GI:21740357
VERSION HTC.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 2616)
AUTHORS Bahr, A., Lauber, J., Mewes, H.W., Well, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
CONSRM The German CDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT This clone (DKFZp434I2411) is available at the RZPD Deutsches
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
This clone (DKFZp434I2411) is available at the RZPD Deutsches
Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp434I2411
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/
Location/Qualifiers
1. .2616
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434I2411"
/db_xref="taxon:9606"
/clone="DKFZp434I2411"
/issue_type="testis"
/clone_id="434 (synonym: htes3). Vector pSPORT1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
/note="hypothetical protein (Rattus norvegicus),
N-terminus truncated"
1. .2616
/gene="DKFZp434I2411"
1. .885
/gene="DKFZp434I2411"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAD39186.1"
/db_xref="GI:21740358"
/db_xref="GO:08NCX8"
/db_xref="UniProt/TREMBL:Q8NCX8"
/translation="RSP1ARBELRLNLFPPPEYGVADGTAVDIFSPGCALEMAV
LEIQNDGTVTEBAIARSHSLSDPMRBSFILCLADAPRPSASLLFHRVLFV
HSLKLAACFIQHOYLMPEVNEKTKMLDLHVALLEPRPRPLQMSYSEVFE
LDKFLDVNRGIVYLMNFATRPGLPRVALPPEVQKXKTRPPPEDETKVIOM
QCNLESEDKARWHLTLLVLEDRILHQLTLYDLLPTDSADLASLVHYGFLNEDDM
KLAFLBESTFLKRGTOA"

ORIGIN
Alignment Scores:
Pred. No.: 6.66e-152 Length: 2616
Score: 1536.00 Matches: 294
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 57.53% Indels: 0
DB: 4 Gaps: 0
US-10-618-941-99 (1-507) x HSM805653 (1-2616)
Qy 214 ArgSerProIleArgAlaGluArgGluGluLeuArgAsnLeuHisPheProProGlu 233
Db 1 CGAAGCCCATCCGCGCTTGAGCGAGAGAACTTGGAACTTGCATTCTTCCCGCAGAG 60
Qy 234 TyrGlyGluValAlaAspGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeu 253
Db 61 TATGAGAGGATGGCGCATGGGACCGCTGTGGACATCTTCTCTTGGGATGTGGCGCTG 120
Qy 254 GluMetAlaValLeuGluIleGlnThrAsnGlyAspThrArgValThrGluAlaIle 273
Db 121 GAGATGGCTGTACTGGAATCCAGAACCAATGGGACACCGGGTCAACAGAGAGGCATT 180
Qy 274 AlaArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysGluLeu 293
Db 181 GCTGGCGCCAGGACCTGCTGAGTGAACCCCAACATGGGAGATTGCTGCTGCTG 240
Qy 294 AlaArgAspProAlaArgAspProSerAlaHisSerLeuLeuPheHisArgValLeuPhe 313
Db 241 GCCCGGAGCCCTGCGCGCCGCGCCCTCTGCCACAGCTCTCTTCCACCGCGTCTTC 300
Qy 314 GluValHisSerLeuTyrLeuLeuAlaHisCysPheIleGlnHisGlnTyrLeuMet 333
Db 301 GAGGTGACCTCGCTGAGAGCTCTGGACGCCACACTGCTTACCGACACAGTACCTCATG 360
Qy 334 ProGluAsnValValGluGluTyrThrValMetAspLeuHisAlaValLeuAlaGlu 353
Db 361 CCGAGAAATGTGGAGGAGAAAGAACCAAGGCCATGGACCGGGTCTTGGCGAG 420
Qy 354 LeuProArgProArgAspProProLeuGlnThrArgTyrSerGluValSerPheMetGlu 373
Db 421 CTTCCCGGCGCCGCGAGGCCCGCCGCTCACTGGCGGTACTCGAAGTCTCTTCAATGAG 480
Qy 374 LeuAspIyrAspLeuGluAspValArgAsnGlyIleTyrProLeuMetAspPheAlaAla 393
Db 481 CTGGACAAATTCGTGAGAGATGTCAGGAATGGAATGATCAACAGATGAATCTTGGAGCC 540
Qy 394 ThrArgProLeuGluTyrLeuProArgValLeuAlaProProProGluValGlnPheAla 413
Db 541 ACTCGACCCCTGGGGCTGCCCGCGTGTCTGGCCCAACCCCGGAGGAGTCCAAAGGCC 600
Qy 414 TyrThrProThrProGluProPheAspSerGluThrArgTyrValIleGlnMetGluCys 433
Db 601 AAGACCCGACCGGAGGCCCTTTGACTGTGAACAGAAAGGATCATTCAGATGCAATGTC 660
Qy 434 AsnLeuGluArgSerGluAspIyrAlaArgTyrHisLeuThrLeuLeuValLeuGlu 453
Db 661 AACCTGGAGAGAGCGAGGACCAAGCGCGCTGGATGTCACTGCTTGGTGTCTGGAA 720
Qy 454 AspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThrAspSerAlaGlnAspLeu 473
Db 721 GACCGGTGACCGGCGAGCTGACCTTCACTGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
Qy 474 AlaSerGluLeuValHisTyrGlyPheLeuHisGluAspArgMetCysValAlaAla 493
Db 781 GCTCGAGAGCTGAGCATAGGCTTCTCTCAAGAGGACGACGAGATGAGGAGGAGGAGGAG 840
Qy 494 PheLeuGluSerThrPheLeuTyrTyrArgGlyThrGlnAla 507
Db 841 TTCTGTGAGAGGACCTTCTCTCAAGTACCGTGGAGGCCAGGCC 882
RESULT 6

AY408904	1601 bp	DNA	linear	GSS 15-DEC-2003
LOCUS				
DEFINITION	Pan troglodytes NRPB gene, VIRTUAL TRANSCRIPT, partial sequence,			
ACCESSION	AY408904			
VERSION	AY408904.1			
KEYWORDS	GSS.			
SOURCE	Pan troglodytes (chimpanzee)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Pan.			
REFERENCE	1 (bases 1 to 1601)			
AUTHORS	Clark,A.G., Gnanowsk,i,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.U., Shinsky,J.U., Adams,M.D. and Cargill,M.			
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios			
JOURNAL	Science 302 (5652), 1960-1963 (2003)			
PUBMED	14671302			
REFERENCE	2 (bases 1 to 1601)			
AUTHORS	Clark,A.G., Gnanowsk,i,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.U., Shinsky,J.U., Adams,M.D. and Cargill,M.			
TITLE	Direct Submission			
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA			
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.			
FEATURES	Location/Qualifiers			
source	1..1601			
gene	/organism="Pan troglodytes"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9598"			
	<1..>1601			
	/gene="NRPB"			
	/locus_tag="HCM3379"			
ORIGIN				
Alignment Scores:				
Pred. No.:	4,51e-152	Length:	1601	
Score:	1534..50	Matches:	305	
Percent Similarity:	73.54%	Conservative:	59	
Best Local Similarity:	61.62%	Mismatches:	102	
Query Match:	57.47%	Indels:	29	
DB:	10	Gaps:	6	
US-10-618-941-99 (1-507) x AY408904 (1-1601)				
QY	17	GIUATGAGUAAGPGLUSERGIUNAPGUSERAEPJLEUGLUGLUSERPROCGIATG	36	
DB	123	GAGGAAGAAAGAAAGTGAAGTGAAGTTCGAGATTTTGAAAGAGTGGCCCTGTGGGCC	182	
QY	37	TTPGINTLYAARGARGLUGLINVALAENGINGLYASMETPROGLYUENGLINSETRPHE	56	
DB	189	TGGCAGAAAGGCGAGAAAGGAGGATCAACGGAATGTACCGAGTATTCACAGTGCATAC	242	
QY	57	LEUAIAMETASPTHTGICUGLUGLYVALGIUVALVATTPASNGLUENHISPEGLYASP	76	
DB	243	CTGGCCATCATGATACAGAGAGAGGTGTAGAGGTGTGTGGAAATGAGGTACAGTTCTT	302	
QY	77	ATGTYVALAIPHEALAHIAHTIAGIUGLYLSIIEGINTHTVALPHEGLUGLNUVALLEU	96	
DB	303	CCCAAGAACTTACAGCTGCAAGGAGGAAAGGTTCGTCTGTGTGTGATATATCATCTCA	362	
QY	97	VALASPIHPROKSNILEVALYLSLEUHSIYSYTYTTPLEASPTHTSERGIUALACYS	116	
DB	363	TTCGAGCACTTATACCTTGTATAGTTTCCAAATATATGGCGTCGACATTAAAGAAACA	422	
QY	117	ALATGVALIIEPHEIIEHTHRTGTYVALSERSEGIYSEIUEYSGINPHEIUEYVS	136	

Db	423	GCACGGGCACTTTTATACAGAAATACATGCACTCTGGAGTGTGAAGCAATTTCTGAG	482
Qy	137	LySThrIuVlyAvAnHlVlyValAMetAspAlaArgAlaTrpIuValGTrpCysThrGln	156
Db	483	AAGACCAAAAAAGAACCCACAGACCATGAAATCAAAAGACATGGAAGCGTTGGTGCACAA	542
Qy	157	ILeuleuSerAlaLeuSerPheLeuHisValCysSerProPoiLeileHisGlyAspLeu	176
Db	543	ATCCTCTCTGGCCCTAAGCTACCTGCACTCTCTGTAACCCCACTCACTCACTGGGAACCTG	602
Qy	177	ThrSerAspThrIlePheIleGlnHisAsnGlyLeuIleLysIleGlySerValTrpHis	196
Db	603	ACCTGTACACACCACTTTCATCCACACACAGCATCTCAAGATGGCTGTGTG-----	656
Qy	197	ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspArgLeuArgSerPro	216
Db	657	-----GGCTGTGACCTATTCACAAATCATC-----GCTCTGTACACTATTCACAAATCATC	680
Qy	217	ILearArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluTrpGlyGlu	236
Db	681	GTGAAGACTTGTGCAGAAAGACAGCAAAATCTCACTTTTGACACCAAGATATGAGAA	740
Qy	237	ValIlaAsp---GlyThrAlaValAspIlePheSerPheGlyMetCysValAlaGluMet	255
Db	741	GTCACTAATGTGCACAAACAGCAAGTGCACATCTACTCTTTGGCAGTGTGCACCTGGAATG	800
Qy	256	AlaValLeuGluIleGlnThrAsnGlyAspThrArg--ValThrGluAlaIleAla	274
Db	801	GCAGTGTGCAGATTTCCAGGCGAAATGGAGAGTCCATATGTGTGCCACCAAGACCATGAC	860
Qy	275	ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysCysLeuAla	294
Db	861	AGTGCACATCCAGCTTTAGAAAGCCATTCAGAGGAGATTCATCAAAAGTCCCTGCAG	920
Qy	295	ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu	314
Db	921	TCTGAGCTGTCTCCGACACCAACAGCCAGAACTTCTGTCCACCAACAGCATTTGTTGAA	980
Qy	315	ValHisSerLeuLysLeuLeuAlaAlaHisCysPheIleGlnHisGlnTrpLeuMetPro	334
Db	981	GTGGCCCTGTCCAAACTCTTGGCGGCCACTGATTTGGGACACCAACATGATCCCA	1040
Qy	335	GluAsnValValGluGluLysThrIleValAMetAspLeuHisValValLeuAlaGluLeu	354
Db	1041	GAGAAACGCTCTGCAGAGAGATCACCAAAACATGTGATCTAGTGTGTGATCTGGCGAAATC	1100
Qy	355	ProArg---ProArgArgProProLeuGlnTrpArgTrpSerGluValSerPheMetGlu	373
Db	1101	CCTGCAGACCAAGGAAGAAACAGTTCAAGCTTTGTATCTTCACGTACCAACAGCTCTGGA	1160
Qy	374	LeuAspLysPheLeuGluAspValArgAsnGlyIleTrpProLeuMetAsnPheAlaAla	393
Db	1161	TTAAGTTAAATCTCTTGAAGATGTCAAGAAATGGATCTATCTCTGCAGACGCTTT-----	1214
Qy	394	ThrArgProLeuGlyLeuProArgValLeuAlaProProArgGluGluVal-----	410
Db	1215	-----GGGCTGTGCTCGGCGCCCGCCAGACGACACAGACAGAGAGAGTACATCACT	1262
Qy	411	-----GlnLysValAlaLysThrProThrProGluProPheAspSerGluThrArgLys	427
Db	1263	GTCTGTCCCCCTCTGTCAAGACTTCCACACTTGAACCACTGAGTGAAGTCTGGCAAG	1322
Qy	428	ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLysValAsnGlyTrpHisIleuThr	447
Db	1323	GTGGTGTGATGAGCTGCATGAACTTATGATGCGGTGAGAGGGAGAGTCAAAACACACTGAG	1388
Qy	448	LeuLeuLeuValLeuGluAspArgLeuHisValArgGlnLeuThrTrpAspLeuLeuProThr	467
Db	1383	CTTCTGTGAGATTTGGAGAGCAAACTGAACCGGACACCTGACSTGTGACSTGTACAAAT	1442
Qy	468	AspSerArgAlaGlnAspLeuAlaSerGlnLeuValHisTrpGlyPheLeuHisGluAspArg	487
Db	1443	GAGAAATATCCCGCAATTTGGCGAGTGTGAGTGTGACSTGGGCTTTCATTTAGTGAAGCTGAC	1502

[illegible]

QY	77	ArgGlySerAlaPheAlaAlaHisLeuGluGlyLeuIleGlnThrValPheGluGlnLeuValLeu	96
Db	409	CGAAGAAGTCAACAGCTGCAGAGAGAGAAAGGGTGTGTGTGTGTGTGAATCTCAATTCAA	468
QY	97	ValAspHisProAlaGlnIleValIleGluLeuHisIleGlyTyrTrpLeuAspTrpSerGluAlaCys	116
Db	469	TTGAGGACATCTTAACATTTGTTTAAGTTTCCACAAATATTTGGCTGACATTAAAGAAACAAG	528
QY	117	AlaArgValIlePheIleThrGluTyrValIserSerGlySerLeuIleGlnPheLeuLys	136
Db	529	GCACGGGCTCATTTTATTCACAGAATTAACATGCATCTGGAGATCTGAAGCAATTTCTGAAG	588
QY	137	LysThrIleLysAsnHisLysValMetAsnAlaArgAlaTPrpLysArgTrpCysThrGln	156
Db	589	AAGACCAAAAAGAACACACAGACATCATATGAAGAAGCATGGAAGCGTTGTCACACAA	648
QY	157	IleLeuSerAlaLeuSerPheLeuHisAlaCysSerProIleIleHisGlyAsnLeu	176
Db	649	ATCCTCTGTGCGCTTAAGCTTACCTGACCTGATCCCTGGATCATTCATGAGAACTGT	708
QY	177	ThrSerAspTrpIlePheIleGlnHisAsnGlyLeuLeuIleLysIleGlySerValTrpHis	196
Db	709	ACCTGTGCACCAATCTTCATTCACACACAGCATCATCAAGATTTGGTCTGTG----	762
QY	197	ArgIlePheSerAsnAlaLeuArgProProThrAlaLeuProAspLeuArgSerPro	216
Db	763	-----GCTCTGACACATCAACATCAT-----	786
QY	217	IleArgAlaGluArgGluGluLeuArgAsnLeuHisPhePheProProGluTyrGlyGlu	236
Db	787	GTGAAGACTTGTCCAGAAAGACACAAAGATCTCACTTCCTTTGGACCAAGATATGAGAA	846
QY	237	ValAlaLeuAsp---GlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMet	255
Db	847	GTCACTAATGTGACACACAGCTGACATCTACTCTCTTTGGCANTGTGCACTGGAGATG	906
QY	256	AlaValLeuGluIleGlnThrAsnGlyAspThrArg---ValThrGluGluAlaIleAla	274
Db	907	GCAGTGTGGAGATTCCAGGGCAATGGAGAGCTCATATGTGGCCACAGAAAGCCATCAGC	966
QY	275	ArgAlaArgHisSerLeuSerAspProAsnMetArgGluPheIleLeuCysCysLeuAla	294
Db	967	AGTCCCATCCAGCTTCTTAAGAGCCATTCACAGAGGAGTTCATCAAAAGTCCCTGAG	1026
QY	295	ArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPheGlu	314
Db	1027	TCTGAGCTGTCCCAACACCAACCAAGAACTTCTGTTCACCCACAGCATTTGTGAA	1086
QY	315	ValHisSerLeuLysLeuLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetPro	334
Db	1087	GTGCCCTGTGCAAACTCTTGCAGGCCACATGTTGGGACACCAACATGATCTCCA	1146
QY	335	GluAsnValValGluGluLysThrThyAlaMetAspLeuHisAlaValLeuAlaGluLeu	354
Db	1147	GAGAAACGCTTGAAGAGATACCAAAAACATGATACTAGTGGCTGATCTGGTAAATC	1206
QY	355	ProArg---ProArgArgProProLeuGlnTrpArgTyrSerGluValSerPheMetGlu	373
Db	1207	CCTGCAGGACACAGAAAGAAACAGTTCAGACTTTGATCTCTCACTCCAGCTCTGGA	1266
QY	374	LeuAspLysPheLeuGluAspValArgAsnGlyLysTyrProLysMetAsnPheAlaAla	393
Db	1267	TTAAGATAAATTCCTTGAAGATGTGAGGAATGGAGATCTATCTCTGCACGCTTT-----	1320
QY	394	ThrArgProLeuGlyLeuProArgValLeuAlaProProProGluGluVal-----	410
Db	1321	-----GGGCTGCTCTGGGCCCCAGAGCCACACAGAGAGAGGTACATCACT	1368
QY	411	-----GlnLysAlaLysThrProThrProGluProPheAspSerGluThrArgLys	427
Db	1369	GTCTGTGCCCCCTGTGCACAGATCTCCACACTTAACAGCTGAGGTGAGTACCTGCAAG	1428

Oy		428	ValliegiImeGInCvAAsnleudlunYrgsergibwArlyVAliaAgTPrhiIseuthr	447
Db		1429	GTGGTCGTAATGCACGTGCACTTAAGTCGCGTGGAAGGAGAGTCAAACAACCATTGACA	1488
Oy		448	LeuleuLeuVallleuGlusprArgIueuIsvArgIuleuThrTyZAvRLeuleuprothr	467
Db		1489	CTTCGCTGAAGTTGGAGAGCAAACTGAACCCGACACTGAGCTGCACTGAATGCCAAAT	1548
Oy		468	AhpSerIagIAnRpLeuAlasergIuleuValniSTyrgIyPheLeuHiGIuAwrApR	487
Db		1549	GAGATATACCCCAGTAGTCGGCGCTGAGTCGGTGCAAGCTGAGGCTTCATTAGAAGGCTAC	1608
Oy		488	ArguetlyvLeuAlAlAPheLeugIuberThrrPheLeuTytyr	502
Db		1609	CAGAGCCGGTTGACTTCTCTGCTGAMAAGACCTTGAACAAGTTC	1653
RESULT 9				
HSM801650				
LOCUS				
DEFINITION	HSM801650	2158 bp	mRNA	linear HTC 22-SEP-2004
ACCESSION	ALJ36682			
VERSION	ALJ36682.1	GI:12052887		
KEYWORDS	HTC.			
SOURCE	Homo sapiens	(human)		
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 2158) Osterweider,B., Obermaier,B., Deuteschbauer,S., Schairp,A., Mewes,H.W., Weill,B., Amlid,C., Osanger,A., Fodor,G., Han,W. and Wiemann,S. The German CDNA Consortium Direct Submission Submitted (22-SEP-2004) MIPS, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY			
CONSRMT				
TITLE	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de/ Sequenced by Medigenomix (Martinied/Germany) within the CDNA sequencing consortium of the German Genome Project. This clone (DKFZp564D1878) is available at the RZPD Deutsches Reisourcenzentrum fuer Genomforschung GmbH in Berlin, Germany. please contact RZPD for ordering: http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp564D1878			
JOURNAL	Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/. location/qualifiers			
COMMENT				
		1..2158		
FEATURES				
SOURCE				
		/organism="Homo sapIens"		
		/mol_type="mRNA"		
		/db_xref="RZPD:DKFZp564D1878"		
		/db_xref="taxon:9606"		
		/clone="DKFZp564D1878"		
		/tissue_type="brain"		
		/clone_id="564 (synonym: hfbr2). Vector pAMP1, host X1-blue; sites NotI + SalI" /dev_stage="fetal"		
		/note="nuclear receptor binding protein"		
		1..2158		
		/gene="DKFZp564D1878"		
		88..1695		
		/gene="DKFZp564D1878"		
		/codon_start=1		
		/product="hypothetical protein"		
		/protein_id="CAB66617.1"		
		/db_xref="GI:12052886"		
		/db_xref="GOA:Q9UHY1"		
		/db_xref="UniProt/TREMBL:Q9UHY1"		
		/ec_number="M5EGSQTVLSGSDDPKVESAPGLTSVPVSTTSASP EEESSEDSRIEESPCGRMKOREEVNONNVEIDSAYLAMPTREEVEVMKVS ERKYVKAKDEEKVAEPNNILQLHELNIVKHFKRADIKENKAYITFTTEWSGSLLQ FLUKTKQHKTMEKAKRMCTOLHSLASTVSHSGCDPIFGNLICDTIFLOHGLIKI GSAVPDPINNHVKTGREEQTLHFPAEYLVSVNVTAVADVIYSFGMCALMAVLIEIG		
gene				
CDS				

NGESSYVQPEALSSAIGLEEDPDLREFIOKLQSPPARPPTARELLFHPALFEVPESL					
LLAAHCTIGHQHMI PENALEETITKXMDTSAYLAELI PABGEPBPVTGLYSQSALRID					
FLBEDVRNGIYPLTAFGPRPOQPQEETSIPVPSGVTKPIPEBPAPERVKVLWQC					
IESVEGVKHLLTLILKLEDKLNRLHSICOLMPENNI PEIIAEVLVOLGFISEADQSRL					
SLEBETINKFPFNARSTLSNAVTVSS"					
 ORIGIN					
Alignment Scores:					
Pred. No.:	7,15e-152	Length:	2158		
Score:	1534.50	Matches:	305		
Percent Similarity:	73.54%	Conservative:	59		
Best Local Similarity:	61.62%	Mismatches:	102		
Query Match:	57.47%	Indels:	29		
DB:	4	Gaps:	6		
US-10-618-941-99 (1-507) x HSM801650 (1-2158)					
Oy	17	GLUARGLUAPRGIUSeRGLUAPRGIUSeRhpIIeLUglUGIUSeRProCysegiYarg	36		
Dd	217	GAGGAAGAAGAAAGTGAAGATGACTGATTTTGAAAGATCGCCCTGTGGGCC	276		
Oy	37	TPrpInLYSAASGXGluGlInValAsnGIngyAsmMetProGlyLeuGlnSerThrPh	56		
Dd	277	TGGCAAGAGAGCGCAGAGAGGTGATCAACGAATGTACCAGGATATTGCACGTGCATAC	336		
Oy	57	LeuAlmeaEprThnglUGluGluValGluValValTPraSnGluLeuNiSPhegLYap	76		
Dd	337	CTGGCCATGATACAGAGAGAGGTGTAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGA	396		
Oy	77	ArgLybaLaPhealaaLahIEglUGluLYslEGInThrValPheGluGlnLeuValLeu	96		
Dd	397	CGCAAGAACTCAAAGCTGCACAGAGAAAAGGTTCGTGCTGTGTGTGTGTGTGTGTCAA	456		
Oy	97	ValasphISproanIlleValLYLeuNiSLyeTYrtTPLeuaSPThRSerGluLaCYs	116		
Dd	457	TTGAGACATCTTAACATGTGAAGTTTACAATAATTTGGCTGACATTTAAAGAAACAAG	516		
Oy	117	AlaaRgaValIIepheIethrgIuTyvAlserSeRgiSerLeuLYsgInPheLeuLYs	136		
Dd	517	GCGAGGGCATTTTATTATCACAGATACATGTCACTGTGGAGTCTGAACAATTTCTGAAG	576		
Oy	137	LysThrLYslYaSNhiSLYvalImecANlaaRgalatRPlyeaRgtRPcyathrgIn	156		
Dd	577	AAGACCAAAAAAGAACCAAGACATGAATGAAGGATGGAAGCGTTGGTGACACACA	636		
Oy	157	IleLeuSERalaleuSerPheLeuNiSaLaCYserProproIleIehIgLYaSNleu	176		
Dd	637	ATCTCTCTGCCCTTAAGCTATCTTCACATCCCTGTGACCCCCCATCATTCATGGGAACCTG	696		
Oy	177	ThisSerAPThrlIepheIegLnHiSaNgLYLeuLielysIIegLYSerValTrpHis	196		
Dd	697	ACCTGTACACCATTTTCATCATCACACACAGCACTCATCAAGATTTGGCTCTGTG-----	750		
Oy	197	ArgIlePheSerANlaLeuayRgProProThralaLeuProaspAsPLeuayRgSerPro	216		
Dd	751	-----GCTCTGACACTATCAACATCAT	774		
Oy	217	ILearGalagLuARGluGluLYLeuayRgANleuNiSPhePheProProGluTYrgLYgu	236		
Dd	775	GTGAAAGCTTGTCAGAGAGACAGACAGATTTTACACTTTTGGACCCAGATATGAGANA	834		
Oy	237	ValAlaaSP---GLYThrrAlaValAspIlePheSerPhegiymetCYeAlaLeuGlnmet	255		
Dd	835	GTCCTATATGTGACACACAGCTGACATCTACCTCTTGGCAGTGTGCACTGGAGATG	894		
Oy	256	AlaValaLeuGluIlegInThraSngLYaSPThraRg--ValThrgIugLuAlaIleAla	274		
Dd	895	GCAGTGCTGGAGATTCAAGGCAATGGAAGTCCCATATGTGGCCAGAGAGCCATCAGC	954		
Oy	275	ArgAlaArghISerLeuSerAPProahmectArgIuPheIIeLUcYsCYeLUaLa	294		
Dd	955	AGTGCCATCCAGCTTTCAGAAAGCCATTACAGAGGAGATTTCATTCAMAGTGCTGTGAG	1014		

QY 295 ArgAspProAlaArgArgProSerAlaHisSerLeuPheHisArgValLeuPheGlu 314
Db 1015 TCTGAGCTCTGCTGCAGACCAACAGCAAGAACTCTGTTCCACCAGCATTTGTTGAA 1074
QY 315 ValHisSerLeuValSerLeuValAlaHisCysPheIleGlnHisGlnTyrLeuMetPro 334
Db 1075 GTGCGCTCGCTCAAACTCCTTGCGGCCCACTGCATTTGTGGAGACCAACCATGATCCCA 1134
QY 335 GluAsnValValIleGluTyrThrTyrValMetAspLeuHisAlaValLeuValIleGluLeu 354
Db 1135 GAAACAGCTCTTAAGAGGATCACCAAAACATGATGATCTGTCCTGATCGCTGAAATC 1194
QY 355 ProArg---ProArgArgProProLeuGlnTyrArgTyrSerGluValSerPheMetGlu 373
Db 1195 CTTGCAGAGACCAAGAAAGAACAGATTCAAGCTTTGATCTTCACTCAAGTCAACAGCTTGGAA 1254
QY 374 LeuAspLysPheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 1255 TTGATTAATATCTCTTGAAGATGTCAGGAATGGGATCTATCTCTGACAGCTTT----- 1308
QY 394 ThrArgProLeuGluTyrProArgValLeuAlaProProProGluGluVal----- 410
Db 1309 -----GGGCTGCCTCGGCCCCAGCACCCACAGCAGAGAGAGGTGACATCACT 1356
QY 411 -----GlnLysAlaLysThrProThrProGluProPheAspSerGluThrArgLys 427
Db 1357 GTTCGTGCCCCCTCTGTCAAGACTCCGACACCTGAACCAAGTAAAGTGAGACTCCCAAG 1416
QY 428 ValIleGlnMetGlnCysAsnLeuGluArgSerGluAspLysAlaArgTyrHisLeuThr 447
Db 1417 GTGTGCTGTATGACAGTGCAACATTGAGTGGTGAAGAGAGGATCAACACCACTGAC 1476
QY 448 LeuLeuLeuValLeuGluAspArgLeuHisArgGlnLeuThrTyrAspLeuLeuProThr 467
Db 1477 CTCTCTCTGAAGTTGGAGGACAAACTGAACCGGCACCTGAGCTGATGACCAAAAT 1536
QY 468 AspSerAlaGlnAspLeuAspSerGluLeuValHisTyrGlyPheLeuHisGluAspArg 487
Db 1537 GAAATATATCCCAAGTTGGGGCTGAGCTGTGACGTGGCTTCAATTGCTAGAGCTGAC 1596
QY 488 ArgMetLysLeuAlaAlaPheLeuGluSerThrPheLeuLysTyr 502
Db 1597 CAAAGCGCGTTGACTTCTCTGTAGAAAGACCTTGAACAAGTTTC 1641
RESULT 10
CR858669 2181 bp mRNA linear HTC 12-NOV-2004
LOCUS Pongo pygmaeus mRNA, cDNA DKF2p459g2032 (from clone DKF2p459g2032).
DEFINITION CR858669
ACCESSION
VERSION CR858669.1 GI:55728271
KEYWORDS
SOURCE HTC
ORGANISM Pongo pygmaeus (orangutan)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Pongo.
REFERENCE 1 (bases 1 to 2181)
AUTHORS Koehler, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
CONSTRM Fob, G., Han, M. and Wiemann, S.
TITLE The German cDNA Consortium
JOURNAL Direct Submission
COMMENT Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764
Neuberberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de,
sequenced by BMFZ (Biomedical Research Center at the
Heinrich-Heine-University, Duesseidorf/Germany) within the cDNA
sequencing consortium of the German Genome Project. This clone
(DKF2p459g2032) is available at the RZPD Deutsches
Resourcenzentrum fuer Genomforschung GmbH in Berlin, Germany.
Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKF2p459g2032
Further information about the clone and the sequencing project is

FEATURES available at <http://mips.gsf.de/projects/cdna/>.
source location/Qualifiers
1..2181
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKF2p459g2032"
/feature_type="cortex"
/clone_lib="459 (synonym: poor1). Vector pSport1_Sfi; host
DH10B; sites SfiI + SfiII"
/dev_stage="adult"
/note="nuclear receptor binding protein (Homo sapiens)"
1..2181
/gene="DKF2p459g2032"
116..1723
/gene="DKF2p459g2032"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAH90881.1"
/db_xref="GI:55728272"
/translation="MSGSSQTVLSGSDPKYSSSSAPGLTVSPPTVSTTSAPSE
EEESDESEILIESPCGRWQREVNQNVPGIDSAVLADTEGVVNEVQFS
ERKNYKIQBEKVAVFDNLIOLEHLNIYKTFHYMADIKENKARVIFITEYSSGLQK
FLKTKKHKTWNEKAMRMCTQILSALSYLHSCDPIHGNLTCDTPIHNGILIKI
GSVADDTNNHYKTCREOKNHPFAPRYGEVTVTTAVDIYSFGMCALFEMVLEIQ
NGESSYVROEALISAIQILEDPLOREPIQKCIQSEPARPRTARELFPRLPEVSEIK
LIAAHCIVGHQMTPENALBEITKMDTSAVLAELIPAGGRPPVQTLVSQSPALEIDK
FLEDRVNGIYPLTAPRPPQPOQBEATVSPVPSVKPTPEPPEVEVTRKVLQCN
IESVEGVGHHTLLTKLEDIKLRLHSICDLMENENIPELIAELVLQGFISEADQSLRT
SLBERTLNKFNPARSSTLNSAIVTVSS"
ORIGIN
Alignment Scores:
Pred. No.: 7.27e-152 Length: 2181
Score: 1534.50 Matches: 305
Percent Similarity: 73.54% Conservative: 59
Best Local Similarity: 61.62% Mismatch: 102
Query Match: 57.47% Indels: 29
DB: 4 Gaps: 6
US-10-618-941-99 (1-507) x CR858669 (1-2181)
QY 17 GluArgGluAspGluSerGluAspGluSerAspIleLeuGluGluSerProCysGlyArg 36
Db 245 GAGGAAGAAGAAAGAGTGAAGATGATGCTGTGAGATTTGGAAAGAGTCGCCATGTGGCCG 304
QY 37 TTPGlnLysArgArgGluGlnValAsnGlnGlyAsnMetProGlyLeuGlnSerThrPhe 56
Db 305 TGGCAGAAAGCGCGAAGAGGTGAATCAACGGAATGTACAGATTTACAGCTGATAC 364
QY 57 LeuAlaMetAspThrGluGluGluValGluValValTTPAsnGluLeuHisAspGlyAsp 76
Db 365 CTGGCCATGATATACAGAGGAAGGTAGAGTGTGTGGAAGAGAGTACAGTTCTTGTGAG 424
QY 77 ArgLysAlaPheAlaAlaHisIleGluGluLysIleGlnThrValPheGluGlnLeuValLeu 96
Db 425 CGCAAGACTAACAGCTGCAGAGGAAAAAGTTTCGTGCTGTGTGTATATCATGATTCAAG 484
QY 97 ValAspHisProAsnIleValLysLeuHisLysTyrTTPLeuAspThrSerGluAlaCys 116
Db 485 TTGAGACATCTTAACATTTGTAAGTTTCACAAAATCTGGCGTGACATTAAAGAAACAAG 544
QY 117 AlaArgValIlePheIleThrGluTyrValSerSerGlySerLeuLysGlnPheLeuLys 136
Db 545 GCCAGGTCATTTTATATCAAGATATACATGTCATCTGGGAGTCTGAAGCAATTTCTGAAG 604
QY 137 LysThrLysLysAsnHisLysValMetAsnAlaAlaGlaTTPLysArgTTPCysThrGln 156
Db 605 AAGACCAAAAGAACCAACAGCATGATGAAAAAGGCATGGAAGCGTTGGTGCACACAA 664
QY 157 IleLysSerAlaLeuSerPheLeuHisAlaCysSerProProIleIleHisGlyAsnLeu 176
Db 665 ATCTCTTCTGCGCTTAAGTACTGCACTCTGTGACCCCCCATCATCATGAGGAACCTG 724

Dd	574	AAAGCAAAAAGACACCAAGACGATGATGAAAAGCATGGAGCGTTGGTCACACAA	633
Qy	157	IIleuSerGAlaleuSerPheIuehIvalAcYseSerProProlIleIleGIYAnIeu	176
Dd	634	ATCGCTCTGGCCCTTAAGCTACCTGCATCTGTGACCCCCCATCATCATCGAGAACTG	693
Qy	177	ThySerAerPThrIlePheIIeGIhIvalenGIYleuIIeYIIeGIYseValTrpHIs	196
Dd	694	ACCTTGACACACCATCTTGCATCCAGACACAGCATCAAGATGGCTGTG-----	747
Qy	197	ArgIIePheSerAvalIaleuAkrProPThrIalAleuProAerPleuAkrSerPro	216
Dd	748	-----GGCTCCGACCATTAACACATAT	771
Qy	217	IIealArgIalAGIuArgIuGIuIeuaIrgAvalIuehIIsPhePheProProlIuTrGIYGIu	236
Dd	772	GTGAAGACTGTGCGGAAAGACGAAAGAAATCTACACTTTTGGACCCAGATHTGAGAA	831
Qy	237	ValIalAer--GIYThrAlaValAerIIePheSerPheGIYmetCysAlaleuGIYmet	255
Dd	832	GTCACTAATGTGACAAACAGACGTGACATCTACTCTTGGCATGTGTGCACCTGGAGATG	891
Qy	256	AlaValIIeGIuIIeGIhIThrAnGIYserPThrArg--ValThnGIuGIuAlIleAla	274
Dd	892	GCAGTGTGGAATTCACAGGCATGAAAGATCTCATATGTGTGCCACAGAAAGCCATCCAGC	951
Qy	275	ArgIalAerGHiserIleuSerAerProAerIleAkrGIuPheIIeIleCysCysIeuaIa	294
Dd	952	AGTGCATATCCAGCTCTTAAGAAACCCA---TTACAGAGTTTCATTCAAAATGCTGTGAG	1008
Qy	295	ArgAerProAlIalAkrArgProSerAlaHiserIleuPhehIvalAkrValIePheGIu	314
Dd	1009	TCTGAGCTGTGCTCGCAGACCAACACCCAGAAATCTGTGCCACCCAGCATTTGTTGAA	1068
Qy	315	ValHiserIleuIYleuIeuaIalAHisCysPheIIeGIhIserGIhIYleuIeuerPro	334
Dd	1069	GTGCCCTGTGCTAACTCTTGGGCCCCACATGTGTGGACACCAACATAGATATCCA	1128
Qy	335	GIuAerValIalGIuGIuIeYThrIYvalAeAerPleuHIsAlaValIeuaIeGIuIeu	354
Dd	1129	GAGAACGCTCTTAAGAGATATCCAAAACATGATTAAGTGGCTTAAGTCTGGCTGAAT	1188
Qy	355	ProAkr--ProAkrArgProProlenGIhItrPargIYSerGIuValIserPheIeGIu	373
Dd	1189	CTTGACAGACCCAGAAAGAACCAATTCAGACTTGTACTCTCAGTACCCAGCTCTGGAA	1248
Qy	374	IeuaerIYpHeIeGIuAerValIalArgenGIYIeIYrProIeuerAerPheAlaIaIa	393
Dd	1249	TTAGTAAATTCCTTAAGATATCCAGAAATGGAACTTACTCTTGACACGCTTT-----	1302
Qy	394	ThraArgProIeugIYleuProAkrValIeuaIaProProGIuGIuValI-----	410
Dd	1303	-----GGGCTGCTTGGCCCCCAGACCCACAGCCAGAGAGAGATGACATCACCT	1350
Qy	411	-----GIhIYAlaIeYThrProThrProGIuProPheAerSerGIuThraIrgIYs	427
Dd	1351	GTGCGTCCCCCTGTGTCAAGATCCCGACACTGAACACAGCGTAGGTGAGATCTGGCAAG	1410
Qy	428	ValIIeGIhImetGIhIcysAnIeGIuIalrgSerGIuAerIYvalAkrTrpHIsIeuthr	447
Dd	1411	GTGGTGTGATGTGACGTGACCACTTAAGTGTGCTGGAGGAGAGTCAAAACACACCTGCA	1470
Qy	448	IeuleuIeuaValIeGIuAerAkrIeuhIIsAkrGIhIleuThrIYrAerPleuIeuerProh	467
Dd	1471	CTTGTGTGAAAGTGTGAGACCAACTGAACCCGGACCTGTAGGTGTGACCTAATGCAAT	1530
Qy	468	AerSerAlaGIuAerPleuAerSerGIuIeuaValIHisIYrGIYpHeIeHIsGIuAerAer	487
Dd	1531	GAGATATATCCCGGAGTGGCGGCTGACGTGGTGGAGCGCTTCAATTAAAGAGGCTGAC	1590
Qy	488	ArgmetIYIeuaIalAkrPheIeGIuSerThrPheIeulYrYr 502	
Dd	1591	CAGACCCGATGTGACTTCTGTGTAAAGAGACTTGAACAAAGTTG 1635	

	RESULT 12
LOCUS AK077440	2181 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 8 days embryo whole body cDNA, RIKEN full-length enriched library, clone:5730409G3 product:SIMILAR TO NUCLEAR RECEPTOR BINDING PROTEIN (HNS7-INTERACTING PROTEIN KINASE) homolog [mus musculus], full insert sequence.	
ACCESSION AK077440	
VERSION AK077440.1	GI:26097350
KEYWORDS HTc; CAP trapper.	
SOURCE Mus musculus (house mouse)	
ORGANISM Mus musculus	
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., TITLE High-efficiency full-length cDNA cloning JOURNAL Meth. Enzymol. 303, 19-44 (1999) PUBMED 10349636	
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., TITLE Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	
JOURNAL PUBMED 11042159	
AUTHORS 3	
Shibata, K., Itoh, M., Aizawa, K., Nagaoke, S., Sasaki, N., Carninci, P., Komno, H., Aktyama, J., Niehl, K., Kitenumi, T., Tashtiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Onate, E., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Mizumatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)	
JOURNAL PUBMED 11076861	
REFERENCE 4	
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
TITLE Functional annotation of a full-length mouse cDNA collection REFERENCE Nature 409, 685-690 (2001)	
AUTHORS 5	
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	
Analytcs of mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)	
JOURNAL REFERENCE 6 (bases 1 to 2181)	
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagahi, T., Harai, A., Hashizume, W., Hayashi, K., Hayatsu, N., Hiramoto, K., Hiraka, T., Hirozane, T., Hoti, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawaji, J., Kohjima, Y., Kondo, S., Komno, H., Koude, M., Koike, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohashi, N., Okazaki, Y., Satto, R., Saitoh, H., Sekai, C., Sakai, K., Sakazume, N., Sano, H., Saeki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sobabe, Y., Tagami, M., Tagawa, A., Takabashi, F., Takaku-Akahira, S., Takeuchi, Y., Tanaka, T., Tomaru, U., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.	
JOURNAL TITLE Direct Submission	
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	
CDDA Library was prepared and sequenced in Mouse Genome COMMENT	

VERSION	CR998804.1	GI:68292689
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 899)	
AUTHORS	Heil, O., Ebert, U., Hennig, S., Henze, S., Radelof, U., Schneider, D. and Korn, B.	
TITLE	Human T-Lymphocytes library	
JOURNAL	Unpublished (2005)	
COMMENT	Contact: Inge Ariart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Email: www.rzpd.de RZPD: RZPDp9017B1314. RZPDLIB: (Human T-Lymphocytes) RZPD LIB No. 9017 http://www.rzpd.de/cgi-bin/products/ret.cgi?libNo=9017 Contact: Inge Ariart RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH Heubnerweg 6, D-14059 Berlin, Germany Tel: +49 30 32639 100 Fax: +49 30 32639 111 www.rzpd.de This clone is available from RZPD. http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPDp9017B1314 contact RZPD (product-support@rzpd.de) for further information. Primer name: qe3.4, Primer sequence: CGGATACAACTTCACACAG. Location/Qualifiers 1. 899 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="RZPDp9017B1314" /tissue_type="T-Lymphocytes" /dev_stage="adult" /lab_host="DH10B" /clone_lib="RZPD no. 9017" /note="Vector: pQE80LSN_cloned; Site_1: SalI; Site_2: NotI; vector: http://www.rzpd.de/info/vectors/pQE80LSN_cloned.pic.shtml ; 1st strand cDNA was prepared from mRNA obtained from human T-Lymphocytes with a NotI - oligo(dT) primer [5' GACGTCTGATGATGCGAGCGCCGCTTTTCTTTTCTTTT 3'] Double-stranded cDNA was ligated to SalI adaptors, digested with NotI and cloned into the NotI and SalI sites of the pQE80LSN_cloned vector"	
ORIGIN		
Alignment Scores:		
Pred. No.:	9,4e-147	Length: 899
Percent:	1481.00	Matches: 290
Best Similarity:	96.31%	Conservative: 1
Best Local Similarity:	97.97%	Mismatches: 4
Query Match:	55.47%	Indels: 2
DB:	7	Gaps: 0
US-10-618-941-99 (1-507) x CR998804 (1-899)		
QY	212	ApPLeuArGSePProIleArGAlaGluArGgUleuArGAsnLeuHlAphePhePro 231
Db	1	GATCTCCGAAGCCCATCCGCGCTAGAGAGAGAACTTCGGAACCTGCACCTCTTCCC 60
QY	232	ProGluTyrGlyGluValAlaAepGlyThrAlaValAapIlePheSerPheGlyMetCys 251
Db	61	CCAGAGTATGGAGAGGTGGCCGATGGACCCGCTGTGTGACATCTTCTCTTGGGATGTGT 120
QY	252	AlaLeuGluMetAlaValLeuGluIleGlnIrrAasnGlyAapThrArGValThrGluGln 271
Db	121	GCGCTGGAGATGCGTGTACTGGAAATCCAGACCAATGGGAGAACCCGGGGGTCAAGAAGAG 180
QY	272	AlaIleAlaArGAlaArGHisSerLeuSerAapProAemMetArGluPheIleLeuGly 291

Db		181	GCCATTGCTGGCGGCAGCACTGTGCTGATGACCCCAACTGTGGGGAGTTCACTTTC	240
Oy	CysleuA1aargAspProAlaArgArgProSerAlaHisserLeuLeuPheH1sarVal	292		311
Db		241	TGCTGTGCCCGGGAGCCTTGCCCGCGGCTCTTGCCACAAGCCTCTTCCACCGCGTG	300
Oy	LeuPhegluValHisserLeuLybleuLeuAlaLHisCysehellleglnHisglnTyr	312		331
Db		301	CTCTTCCAAGGTGCACTGTGTGAAGCTCTCTTGCGACGCCACGTCTTAACCAGCACAGTAC	360
Oy	LeuMetProgluLeuAnValIglugluTyrThrLyalaMetAspleuH1saIvalLeu	332		351
Db		361	CTCATGCTGAAGAATGTGTGGAGGAAACAAGGCCAATGACCTTGACGCGGCTTG	420
Oy	AlaaglueuProArgrProArGrArgProProLeuglnTrpargrTyrsarglualSerPhe	352		371
Db		421	GGGAGGCTTCCC GCCGCCCCGACAGGCCCCCGCTGTGAGTGGGGTACTCGAAGTCTCTTC	480
Oy	MetgluLeuApLyPheLeugluuAspValArganglylleTyrProLeuMetAsnpe	372		391
Db		481	ATGAGAGCTGGACAATAATCTCGAGAGATGTCAAGATGAATCTAACCACTGATGAATCTT	540
Oy	AlaAlaThrArgProLeuglyLeuProArgValLeuAlaPropProProgluIguValGln	392		411
Db		541	GCACCCACTGACCCCTCGGGGCTGCCCCGTGTGTGCCCCACCCCGGAGAGAGTCAA	600
Oy	LySaAlaYsThrProThrProgluProPheAspSerGlunThrArgLySValIIeglnMet	412		431
Db		601	AAGGCCAAGACCCCGAGCCAGAGCCCTTTGACCTGTAGACACGAAAGTCATCAATG	660
Oy	GlnCyasauleugluArgsergluAspLyAlaArgTprHIsleuthrLeuLeuVal	432		451
Db		661	CAGGCAACTCGAGAGAAACGAGAACAAAGCGCGCTGACATCTCACTGCTTGANG	720
Oy	LeugluAspArgLeuH1sarArgInleuThrTyrAspleuLeuProThrAspSerAlaGln	452		471
Db		721	CTGGAMACCGCGCGACCGGACGTCACTTACGACCTGCCCMACGACAGCGCCAG	780
Oy	AspleuAlasergluLeuValHisTyrGlyPheLeuH1egluaAspArg-MetLyble	472		491
Db		781	GACCTCGCTCGAGCTCGTGCCTANTGGCTCTTCCACGANAGACGCCGAAATGAANCT	840
Oy	uAlaAlaPheLeugluSerThrpheLeuLyTyArgGlyThrcIn	491		506
Db		841	GGCGGNTTCTCTGAGAGCACTTCTC-AAGTACCGTGGGACCAAG	885
RESULT 14				
LOCUS	BX410189	927 bp	mRNA	linear EST 15-MAY-2003
DEFINITION	BX410189 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone			
ACCESSION	CS0DN002YI09	5-PRIME,	mRNA sequence.	
VERSION	BX410189			
KEYWORDS	BX410189.1 GI:30768891			
SOURCE	EST.			
ORGANISM	Homo sapiens (human)			
	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
	Homnidae; Homo.			
	1 (bases 1 to 927)			
REFERENCE	Ll,M.B., Gruber,C., Jeesee,J. and Polayes,D.			
AUTHORS	Full-length cDNA libraries and normalization			
TITLE	Unpublished (2001)			
JOURNAL	Contact: Genoscope			
COMMENT	Genoscope - Centre National de Sequencage 2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE Email: betre@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Feng liang Email : fliang@life tech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradise Avenue Genoscope sequence ID : CS0BAG070ZB10_CS06698_2. Location/Qualifiers			
FEATURES				

Source

1. .927
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN002Y109"
/class_type="ADULT BRAIN"
/dev_stage="adult"
/clone_id="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6, 1st strand cDNA was primed with a NotI-oligo(dT) primer, five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

Alignment Scores:

Pred. No.: 2.06e-138 Length: 927
Score: 1403.00 Matches: 275
Percent Similarity: 94.24% Conservative: 3
Best Local Similarity: 93.22% Mismatches: 15
Query Match: 52.55% Indels: 2
DB: Gaps: 0

US-10-618-941-99 (1-507) x BX410189 (1-927)

Qy 214 ArgSerProIleArgAlaGluArgGluLeuArgAsnLeuHisPheProProGlu 233
Db 3 CGAAGCCCATCCCGCTGACGAGAGAACTTCGAACCTGCACTTCTCCCGCAGAG 62

Qy 234 TyGtGluValAlaAerGlyThrAlaValAspIlePheSerPheGlyMetCysAlaLeu 253
Db 63 TATGAGAGGTGGCGCATGGGACCGCTGTGGACATCTTCTCTTGGAGTGTGGCTG 122

Qy 254 GluMetAlaValLeuGluIleGlnThrAsnGlyAspThrArgValThrGluAlaIle 273
Db 123 GAGATGGCTGTACTCGAAATCCAAACCAATGGGACACCGGGTCAACAGAGGCGCAT 182

Qy 274 AlaArgAlaArgHisSerLeuSerAspProAsnMetArgIleLeuLeuCysValLeu 293
Db 183 GCTCGCCCGCAGGCACTGCTGAGGACCCCAACATGGGAGTTCACTTCTGCTGCTG 242

Qy 294 AlaArgAspProAlaArgArgProSerAlaHisSerLeuLeuPheHisArgValLeuPhe 313
Db 243 GCCCGGAGCCCTGGCCGCGCTCTGCCACAGCTCTCTTCCACCGCGTCTTCT 302

Qy 314 GluValHisSerLeuLeuLeuAlaHisCysPheIleGlnHisGlnIleGlnLeuMet 333
Db 303 GAGGTGCACTCGTGAAGCTCTGCGCAGCCCACTGCTTCACTCCAGCACAGTACCTCATG 362

Qy 334 ProGluAsnValAlaGluGluValThrAlaMetAspLeuHisAlaValLeuAlaGlu 353
Db 363 CCGAGATGTGTGGAGAGAGAACCAAGCCCATGTGACCGGGTCTTGGCGGAG 422

Qy 354 LeuProArgProArgArgProProLeuGlnIleArgTyrSerGluValSerPheMetGlu 373
Db 423 CTTCGCCGCGCCCGCAGGCCCCCGCTGCACTGGCGGTACTCGAAGTCTCTTCAATGAG 482

Qy 374 LeuAspIlePheLeuGluAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaAla 393
Db 483 CTGAGCAAAATTCCTGAGATGTCAAGAAATGGAATCAACAGTGAATGAACTTTGGAGCC 542

Qy 394 ThrArgProLeuGluLeuProArgValLeuAlaProProProGluGluValGlnIleValAla 413
Db 543 ACTCGAACCTGTGGGCTGCGCCGCTGTGCTGCCCAACCGGAGGAGGTCCAAAGGCC 602

Qy 414 LysThrProThrProGluProPheAspSerGlyThrArgTyrValIleGlnMetGlnCys 433
Db 603 AAGACCCCGACGCCAGAGCCCTTGAATCTGAGACCGAAGGTCAATCCAGATGCACTGCC 662

Qy 434 AsnLeuGluArgSerGluAspValAlaArgTyrHisLeuThrLeuLeuValLeuGlu 453
Db 663 AACCTGGAGAGAGCGAGCAAGGCGCGCTGGCATCTCACTGCTTCTGAGCTTGA 722

Qy 454 -AspArgLeuHisArgGlnLeuThrTyrAspLeuProThrAspSerAlaGlnAspLe 473
Db 723 AGACGGCGCTCACCCGANNAGTGAACCTTACGAACTGTCCACCGGACAGGCCCCAGACT 782

Qy 473 ValAspGluLeuValHisTyrGlyPheLeuHisGluAspAspArgMetLeuAlaIle 493
Db 783 CGCCTTGAGACTCGGGGACATATGAGCTTCTCCACAGAAACCCGATGAAGCTGGCCCG 842

Qy 493 AspLeuGlu-SerThrPheLeuLysTyrArgGlyThrGln 506
Db 843 CTTTCTGAAAGGCCCTTCTTAAGTCCCGGAGGAGCCAG 883

RESULT 15
B1645887/c 825 bp mRNA linear EST 12-SEP-2001
LOCUS
DEFINITION
60327555F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5315916 5',
mRNA sequence.
ACCESSION
B1645887
VERSION
B1645887.1 GI:15560123
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Muridae; Murinae; Mus.
REFERENCE
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strauberg, Ph.D.
Email: cgaspe-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Place: LLNL1800 row: n column: 13
High quality sequence stop: 808.
Location/Qualifiers
1. 825
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5315916"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/clone_id="NCI_CGAP_Mam3"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI,
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

FEATURES
source

Alignment Scores:

Pred. No.: 6.21e-120 Length: 825
Score: 1229.00 Matches: 239
Percent Similarity: 92.86% Conservative: 8
Best Local Similarity: 89.85% Mismatches: 11
Query Match: 46.03% Indels: 8
DB: Gaps: 1

US-10-618-941-99 (1-507) x B1645887 (1-825)

Qy 140 LysAsnHisLysAlaMetAsnAlaArgAlaTyrGlyArgTyrCysThrGlnIleLeuSer 159
Db 776 CGACGAGTGTGTCGCTCCGCTGCGCAGGCCCCCTGGAAACGCTGTGATCGCAGATCCGTCT 717

```
QY 160 AlaLeuSer-PheLeuHisAla-CysSerProProIleIleHisGluAsnLeuThrSerA 179
   |||||
Db 716 GCACTCAGTTTTTTGGACGCCCCCATCATCGGAGACCTGACGAGCG 657
   |||||
QY 179 sPTThrIlePheIleGlnHisAsnGlyLeuIleIleValIleGlySerValTyrPHisArgIleP 199
   |||||
Db 656 ACACCATCTTCACTTACGACCAATGGCCTCATCAAGATCGGCTCGGTGTGTACCGCATCT 597
   |||||
QY 199 heseSerAsnAlaLeuArgProProThrAlaLeuProAspAspLeuArgSerProIleArgA 219
   |||||
Db 596 TCTCCCAT-----GCACCTTCTGATGATCTCCGAGGCCATTATTCGAG 555
   |||||
QY 219 laGluArgGluGluLeuArgAsnLeuHisPhePheProGluTyrGluGluValAlaA 239
   |||||
Db 554 CTGAACGGAGAACTCCGAACTCGAATCTTTCCACCAAGATATGGCAAGTCAATG 495
   |||||
QY 239 sPGLyThrAlaValAspIlePheSerPheGlyMetCysAlaLeuGluMetAlaValLeug 259
   |||||
Db 494 ATGGGACTGCTGGGACATCTTCTCTCGGAGTGTGTGCACTGGAGATGGCTGTACTCG 435
   |||||
QY 259 IuIleGlnThrAsnGlyAspThrArgValThrGluGluAlaIleAlaArgAlaArgHis 279
   |||||
Db 434 AGATCCAGGCCAACGGGGATACAGAGTCACAGAAAGGCCATCGCTCGAGCCAGGCACCT 375
   |||||
QY 279 erLeuSerAspProAspMetArgGluPheIleLeuGlyCysGluAlaArgAspProAlaA 299
   |||||
Db 374 CACTGAGTGACCCCAACATCGGGAAATCATCTCTGCTGCTGGCCCGGACCTGCCC 315
   |||||
QY 299 rGARProSerAlaHisSerLeuLeuPheHisArgValLeuPheGluValHisSerLeuL 319
   |||||
Db 314 GCCGACCTTCAGCCCAACACTCTCTTCCACCGAGTGTCTTTGAGTGTCACTCGCTGA 255
   |||||
QY 319 yAlaLeuAlaAlaHisCysPheIleGlnHisGlnTyrLeuMetProGluAsnValValG 339
   |||||
Db 254 AGCTGCTGGCAGCTCACTGCTTCATCCAGCACAGTACTCATGCTGAGAAATGTGTAG 195
   |||||
QY 339 luGluIleThrIleValIleMetAspLeuHisAlaValLeuAlaGluLeuProArgProArgA 359
   |||||
Db 194 AGGAAAGAGACCAAGGCCATGACCTTCATGCACTTTGGCTGAGATGCCGAGCCCCCATG 135
   |||||
QY 359 rGProProLeuGlnITrPAArgTyrSerGluValSerPheMetGluLeuAspLysPheLeug 379
   |||||
Db 134 GACCCCAATGCACTGGCGGTACTCAGAGGTCTCTTCTTGAGCTGGACAAATTCCTAG 75
   |||||
QY 379 luAspValArgAsnGlyIleTyrProLeuMetAsnPheAlaIleThrArgProLeuGlyL 399
   |||||
Db 74 AGGATGTCAGAGAACGGGATCTATCACTGATGAACCTTTCGGCTGCTCGGCCCTTGGGGC 15
   |||||
QY 399 euProArgValLeu 403
   |||||
Db 14 TTCCCCGTGTGTG 1
```

Search completed: January 14, 2006, 00:46:08
Job time : 4643 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using SW model

Run on: January 12, 2006, 11:18:45 : Search time 24 Seconds
(without alignments)
1746.524 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670

Sequence: 1 MAPEPAPRRARERERERED.....RKKLAFLSTFLKRYGTA 507

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/prodata/1/1aa/5.COMB.pep.*
2: /cgn2_6/prodata/1/1aa/6.COMB.pep.*
3: /cgn2_6/prodata/1/1aa/H.COMB.pep.*
4: /cgn2_6/prodata/1/1aa/PCTUS.COMB.pep.*
5: /cgn2_6/prodata/1/1aa/R.COMB.pep.*
6: /cgn2_6/prodata/1/1aa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1580	59.2	326	2	US-09-345-473E-10 Sequence 10, Appl
2	1580	59.2	326	2	US-09-862-027-10 Sequence 10, Appl
3	1534.5	57.5	535	1	US-08-933-750C-20 Sequence 20, Appl
4	1534.5	57.5	535	2	US-09-234-613-20 Sequence 20, Appl
5	1534.5	57.5	535	2	US-09-949-016-7299 Sequence 10, Appl
6	1508.5	56.5	535	2	US-09-023-942A-10 Sequence 10, Appl
7	543	20.3	231	2	US-09-345-473E-36 Sequence 36, Appl
8	543	20.3	231	2	US-09-862-027-36 Sequence 36, Appl
9	415.5	15.6	461	2	US-09-345-473E-35 Sequence 35, Appl
10	415.5	15.6	461	2	US-09-862-027-35 Sequence 35, Appl
11	372	13.9	516	2	US-09-345-473E-34 Sequence 34, Appl
12	372	13.9	516	2	US-09-862-027-34 Sequence 34, Appl
13	368.5	13.8	1601	2	US-09-345-473E-40 Sequence 40, Appl
14	368.5	13.8	1601	2	US-09-862-027-40 Sequence 40, Appl
15	364.5	13.7	110	2	US-09-270-767-31733 Sequence 31, Appl
16	364.5	13.7	110	2	US-09-854-856-60 Sequence 60, Appl
17	359.5	13.5	741	2	US-10-010-720-60 Sequence 60, Appl
18	359.5	13.5	741	2	US-09-854-856-60 Sequence 60, Appl
19	359.5	13.5	769	2	US-10-010-720-44 Sequence 44, Appl
20	359.5	13.5	769	2	US-09-854-856-28 Sequence 28, Appl
21	359.5	13.5	801	2	US-10-010-720-28 Sequence 28, Appl
22	359.5	13.5	829	2	US-09-854-856-12 Sequence 12, Appl
23	359.5	13.5	829	2	US-10-010-720-12 Sequence 12, Appl
24	359.5	13.5	894	2	US-09-854-856-54 Sequence 54, Appl
25	359.5	13.5	894	2	US-10-010-720-54 Sequence 54, Appl
26	359.5	13.5	922	2	US-09-854-856-38 Sequence 38, Appl
27	359.5	13.5	922	2	US-09-854-856-38 Sequence 38, Appl

28	359.5	13.5	922	2	US-10-010-720-38 Sequence 38, Appl
29	359.5	13.5	954	2	US-09-854-856-22 Sequence 22, Appl
30	359.5	13.5	954	2	US-10-010-720-22 Sequence 22, Appl
31	359.5	13.5	982	2	US-09-854-856-6 Sequence 6, Appl
32	359.5	13.5	982	2	US-10-010-720-6 Sequence 6, Appl
33	359.5	13.5	1911	2	US-09-854-856-64 Sequence 64, Appl
34	359.5	13.5	1911	2	US-10-010-720-64 Sequence 64, Appl
35	359.5	13.5	1939	2	US-09-854-856-48 Sequence 48, Appl
36	359.5	13.5	1939	2	US-10-010-720-48 Sequence 48, Appl
37	359.5	13.5	1971	2	US-09-854-856-32 Sequence 32, Appl
38	359.5	13.5	1971	2	US-10-010-720-32 Sequence 32, Appl
39	359.5	13.5	1999	2	US-09-854-856-16 Sequence 16, Appl
40	359.5	13.5	1999	2	US-10-010-720-16 Sequence 16, Appl
41	359.5	13.5	2004	2	US-09-854-856-58 Sequence 58, Appl
42	359.5	13.5	2004	2	US-10-010-720-58 Sequence 58, Appl
43	359.5	13.5	2032	2	US-09-854-856-42 Sequence 42, Appl
44	359.5	13.5	2032	2	US-10-010-720-42 Sequence 42, Appl
45	359.5	13.5	2048	2	US-09-854-856-62 Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-09-345-473E-10
Sequence 10, Application US/09345473E
Patent No. 6558903
GENERAL INFORMATION:
APPLICANT: Hodge, Martin
TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
FILE REFERENCE: 35800/183781
CURRENT APPLICATION NUMBER: US/09/345,473E
NUMBER OF SEQ ID NOS: 62
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 326
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(326)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-345-473E-10

Query Match 59.2%; Score 1580; DB 2; Length 326;
Best Local Similarity 97.4%; Pred. No. 1.2e-145;
Matches 301; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY	1	MAPEPAPRRAREREREREDSESDIIEBSPCGRWQRRQVNOGNMPLQSTFLAND 60	
DB	12	MAPEPAPRRAREREREREDSESDIIEBSPCGRWQRRQVNOGNMPLQSTFLAND 71	
QY	61	TEEGEVVWNEHLPGRKFAFAHEEKIQVFQOLVVDHPNVKJHKWLTDSACARYI 120	
DB	72	TEEGEVVWNEHLPGRKFAFAHEEKIQVFQOLVVDHPNVKJHKWLTDSACARYI 131	
QY	121	FTTEVSSGSLKQFLKTKGNKHAMARAKMKTQIISALSFLLACSPPIIHGULTSDT 180	
DB	132	FTTEVSSGSLKQFLKTKGNKHAMARAKMKTQIISALSFLLACSPPIIHGULTSDT 191	
QY	181	FTIHNGLIKISVWHRIFSNALRPPTALPDLRSPIRAREBELNHLFFPEYGEVADG 240	
DB	192	FTIHNGLIKISVWHRIFSN-----ALPDLRSPIRAREBELNHLFFPEYGEVADG 245	
QY	241	TAVDIPSGMGLMAVLEIQNGDTRVTEAIAARSHLSPNMBEFTLCLADPAR 300	
DB	246	TAVDIPSGMGLMAVLEIQNGDTRVTEAIAARSHLSDPNMBEFTLCLADPAR 305	
QY	301	PSVHSLFLH 309	
DB	306	PSVHSLFLH 314	

```
RESULT 2
US-09-862-027-10
; Sequence 10, Application US/09862027
; Patent No. 6858418
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(326)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-862-027-10

Query Match          59.2%; Score 1580; DB 2; Length 326;
Best Local Similarity 97.4%; Pred. No. 1,2e-145;
Matches 301; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 MAPEPAPRRAREREREDESEDESDIIESEPCGWQKREOVNQGNNPGLQSTFLAND 60
    |||||
DB 12 MAPEPAPRRAREREREDESEDESDIIESEPCGWQKREOVNQGNNPGLQSTFLAND 71
    |||||

QY 61 TEEGVVNNELHFGDRKAPAAHEKIQTFEQLVVDHPNIVKLHKYLDTSEACARYI 120
    |||||
DB 72 TEEGVVNNELHFGDRKAPAAHEKIQTFEQLVVDHPNIVKLHKYLDTSEACARYI 131
    |||||

QY 121 FITEVSSGSLKQFLKTKKNHAKAMARAKMCTQIISALSFLLHACSPPIIHGNTSDT 180
    |||||
DB 132 FITEVSSGSLKQFLKTKKNHAKAMARAKMCTQIISALSFLLHACSPPIIHGNTSDT 191
    |||||

QY 181 IFIGHNLIKIGSVWHIRISNAPPTALPDDLRSPIRARREBLRNLFPPPEYGEVADG 240
    |||||
DB 192 IFIGHNLIKIGSVWHIRISN-----ALPDLRSPIRARREBLRNLFPPPEYGEVADG 245
    |||||

QY 241 TAVDIFSPGMCALEMVLEIQTNGDTRVTEEAIAARHSLSDPNMREFTLCCLADPAR 300
    |||||
DB 246 TAVDIFSPGMCALEMVLEIQTNGDTRVTEEAIAARHSLSDPNMREFTLCCLADPAR 305
    |||||

QY 301 PSAHSLFLH 309
    |||||
DB 306 PSAHSLFLH 314
    |||||

RESULT 3
US-08-933-750C-20
; Sequence 20, Application US/08933750C
; Patent No. 5932442
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Bandman, Olga
; APPLICANT: Shah, Purvi
; APPLICANT: Au-Young, Janice
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN REGULATORY MOLECULES
; NUMBER OF SEQUENCES: 98
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
```

```
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ FILING DATE: September 23, 1997
/ CLASSIFICATION: 536
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0356 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 20:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 535 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: BRAITUT03
/ CLONE: 864683
US-08-933-750C-20

Query Match          57.5%; Score 1534.5; DB 1; Length 535;
Best Local Similarity 61.6%; Pred. No. 7e-141;
Matches 305; Conservative 59; Mismatches 102; Indels 29; Gaps 6;

QY 17 EREDESESDIIESEPCGWQKREOVNQGNNPGLQSTFLANDTEEGVVNNELHFGD 76
    |||||
DB 44 EEESESESEIIESEPCGWQKREOVNQNVPIDAVIAMDTEEGVVNNELHFGD 103
    |||||

QY 77 RKAFAAHEKIQTFEQLVVDHPNIVKLHKYLDTSEACARYIFITEVSSGSLKQFLK 136
    |||||
DB 104 RKNYLGQEKRAVVDNLIQLEHNLIVKFKYMDIKENKARVIFITEVSSGSLKQFLK 163
    |||||

QY 137 KTKKNHAKAMARAKMCTQIISALSFLLHACSPPIIHGNTSDTIFIGHNLIKIGSVH 196
    |||||
DB 164 KTKKNHAKAMARAKMCTQIISALSYLHSCDPPIIHGNTCDTIFIGHNLIKIGSV-- 221
    |||||

QY 197 RIFSNALPPTALPDDLRSPIRARREBLRNLFPPPEYGEVAD--TAVDIFSPGMCAL 255
    |||||
DB 222 -----APDTINNHVKTCEQKNLHFAPEYGEVNTTAVDIISFGMCAL 269
    |||||

QY 256 AVLEIQTNGDTR-VTEEAIAARHSLSDPNMREFTLCCLADPARPSAHSILFTRVLFE 314
    |||||
DB 270 AVLEIQTNGESSYVQEAISALQLLEDPQLREFIQKLGQSEPARPFAABELFHPALFE 329
    |||||

QY 315 VHSKLAAHGFIOHYLMPENVVEEKTAMDLAVALER-PRRPLQMRYSVSFME 373
    |||||
DB 330 VPSLGLAAHCTGVGHQWIPENALAEITKMDTSVALAIEIPAGPBRBVQTLYSQSPAL 389
    |||||

QY 374 LDKFLDYRNGIYPLMNAATRPGLPRVLAPRPEV-----QAKTPTPEPSPSETRK 427
    |||||
DB 390 LDKFLDYRNGIYPLTAF-----GLPRPQOQEEVTSVVPVPSVKTPTEPAVERRK 443
    |||||

QY 428 VIWQCNLERSEDKARWHLTLVLVLEDRILHQLTYDLIPTSAODLASLAVHYGFLHDD 487
    |||||
DB 444 VIWQCNLESVEEGKHLTLILKLEDKLNLHLSGDLMPNNINIPBLAELVQLGISSEAD 503
    |||||

QY 488 RMKLAFLFESTFLKY 502
    |||||
DB 504 QSRLLTSLEETLANKF 518
    |||||
```


Db 357 VPSLKLAAHCVGHQHMIPENALBEITGNMDSAVLAIEIPAGPGBEPVQLTYSQSPALE 416
Qy 374 LDKFLEDVRNGIYPLMNFATRLPLGLPRVLAPPEV-----QKAKTTPPEPDSERK 427
Db 417 LDKFLEDVRNGIYPLTAF-----GLPRPOQPQOEVTSPVPPSVKTPPEPAVEYERK 470
Qy 428 VIOMQCNLERSEDKARWHLTLVLVEDRLHROLTYDLPTDSAODLASELYHGFLEHDD 487
Db 471 VILMOQCNIESVEBGVGHKHTLLKLKEDKLNHLSCDLMNENIPELAAELVOLGFISEAD 530
Qy 488 RMKLAFLFESTFLKY 502
Db 531 QSRLTSLLEETLNKF 545
RESULT 6
US-09-023-942A-10
; Sequence 10, Application US/09023942A
; Patent No. 6479274
; GENERAL INFORMATION:
; APPLICANT: (US only) ANTALIS Toni Marie and HOOPER John David
; TITLE OF INVENTION: NOVEL MOLECULES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 GARDEN CITY PLAZA
; CITY: GARDEN CITY
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,942A
; FILING DATE: 13-FEB-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P05101/97
; FILING DATE: 13-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: P04022/97
; FILING DATE: 18-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: International PCT Application
; FILING DATE: 13-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGIGLIO, FRANK S
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 11168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742 4343
; TELEFAX: (516) 742 4366
; TELEEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-023-942A-10
Query Match 56.5%; Score 1508.5; DB 2; Length 535;
Best Local Similarity 60.8%; Pred. No. 2.4e-138;
Matches 301; Conservative 60; Mismatches 105; Indels 29; Gaps 6;
Qy 17 EDEDESDSDIIEBSPCGMRQKRRQVNGMNPGLQSTFLAMDTEBEGVVMNLEHFGD 76
Db 44 EEBEBSDESEILBESPCGRWQRRREVQNRVPGIDSAYLAMDTEBEGVVMNLEHFGD 103

Qy 77 RKAFAHEEKIQTFVEQDLVDHPNIVLKHKMYLDTSEACARVITFEVSGSLKQELK 136
Db 104 RKNYLOGEKVCAPFDNILOLEHINIVKFKYMDIKENKARVIFITGMSSGSLKQELK 163
Qy 137 KTKGNKMANARAWKRWCTQILSALSFHACSPPIIHGNLTSDTIFIQHNGLIKIGSVH 196
Db 164 KTKGNHQTMMERKAWKRWCTQILSALSYLHSCDPIIHGNLTCDTIFIQHNGLIKIGSVH 221
Qy 197 RIFSAARPPALPDRLSPFRAREELRNHAPPEGEVAD-CTAVDIFSGCALDM 255
Db 222 -----APDTINNHVKTCEQKNLHFAPBEGEVNTVTAVDIYSFGCALGM 269
Qy 256 AVLEIQTNGDTR-VTEBAIARARHSLSDPNRREFILCCIADPARPARSHSLFFRVLFE 314
Db 270 AVLEIQNGESSYVPOEAISSAIOLEBPLQREFIOKLOSEPAPRRPFARELLFFPALFE 329
Qy 315 VHSKLKLAACFIOHQYIMPENVEEKTAVDLHVALELPR-PRRPLQWRYSEVSFME 373
Db 330 VPSLKLAAHCVGHQHMIPENALBEITGNMDSAVLAIEIPAGPGBEPVQLTYSQSPALE 389
Qy 374 LDKFLEDVRNGIYPLMNFATRLPLGLPRVLAPPEV-----QKAKTTPPEPDSERK 427
Db 390 LDKFLEDVRNGIYPLTAF-----GLPRPOQPQOEVTSPVPPSVKTPPEPAVEYERK 443
Qy 428 VIOMQCNLERSEDKARWHLTLVLVEDRLHROLTYDLPTDSAODLASELYHGFLEHDD 487
Db 444 VILMOQCNIESVEBGVGHKHTLLKLKEDKLNHLSCDLMNENIPELAAELVOLGFISEAD 503
Qy 488 RMKLAFLFESTFLKY 502
Db 504 QSRLTSLLEETLNKF 518

RESULT 7
US-09-345-473E-36
; Sequence 36, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin
; TITLE OF INVENTION: No. 6558903el kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 231
; TYPE: PRT
; ORGANISM: C. elegans
; US-09-345-473E-36
Query Match 20.3%; Score 543; DB 2; Length 231;
Best Local Similarity 56.3%; Pred. No. 1.2e-44;
Matches 98; Conservative 31; Mismatches 45; Indels 0; Gaps 0;
Qy 19 EDESESDSDIIEBSPCGMRQKRRQVNGMNPGLQSTFLAMDTEBEGVVMNLEHFGRK 78
Db 23 DSDADGAEIILBESPDKMSKRRQVDVPGIDVAIYLANDNETGNEVVMNEVQFSERK 82
Qy 79 AFAAHEEKIQTFVEQDLVDHPNIVLKHKMYLDTSEACARVIFITEVYSSGSLKQFLKKT 138
Db 83 NFAAGEEKINAVFDNLQVLVHTNLVFKFKYMTDSKSKPRIFIFTEWSSGSMAPLQRT 142
Qy 139 KKNHKAANARAWKRWCTQILSALSLFACSPPIIHGNLTSDTIFIQHNGLIKIG 192
Db 143 RKAGSLSLKAWKRWCTQILSALNVLHSDDEPIIHGNLTCTVTFIQNGLIKIG 196
RESULT 8
US-09-862-027-36
; Sequence 36, Application US/09862027
; Patent No. 6658418
; GENERAL INFORMATION:

```

; APPLICANT: Hodges, Martin R.
; TITLE OF INVENTION: No. 6558418el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 231
; TYPE: PR
; ORGANISM: C. elegans
US-09-862-027-36

Query Match      20.3%; Score 543; DB 2; Length 231;
Best Local Similarity 56.3%; Pred. No. 1.2e-44;
Matches 98; Conservative 31; Mismatches 45; Indels 0; Gaps 0;

QY 19 EDESEBDSILFESSPCGRMQRREOVNMGPGLOSTFLAMPDEBEGVYVNMELHRODRK 78
DB 23 DSDADGAEELFESSPDGRMSKRREOVNQRDPGIDVAYLMDNETGVNNEVQFSERK 82
QY 79 AFAAHEKIQTVFEQVLVDHPNIVKLAHYMLDTSACARVIFITEXVSSGSLKQPLKKT 138
DB 83 NFRAGSEKINAVDNLQVLHTLVKFKHKTWDSKSKRPRIITITEXVSSGSLKQPLKKT 142
QY 139 KKNHKNMAMARWKCOTQIISALSFHACSPPIIHGNTLSDTIFIOHNGIKIG 192
DB 143 RKAQSSLSIKVWKMTQIISALNLYLHSSDPPIIHGNTLSDTIFIOHNGIKIG 196

RESULT 9
US-09-345-473E-35
; Sequence 35, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodges, Martin
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 461
; TYPE: PR
; ORGANISM: C. elegans
US-09-345-473E-35

Query Match      15.6%; Score 415.5; DB 2; Length 461;
Best Local Similarity 29.9%; Pred. No. 8.9e-32;
Matches 116; Conservative 57; Mismatches 92; Indels 123; Gaps 13;

QY 210 PDLRSPFAERBELNLHFFPEYGEVADGT-AVDIFSGMCALEMV---LEIQNGD 265
DB 14 PDAINHNVKTCRNMKMYHIAPEYENNTLISADIVSFGICSLIAVIGLSGCGNS 73
QY 266 TR--VTEEAIAARSHSLSDPNKREFLLCCLARDPAPRPSAHSLLFRVLFEVHSLKLLAA 323
DB 74 SEGPVTEVDIEKAIKRLLEDPMQODFIRQCLRKDPAPRPSARELLFQIILFEVHSLKLLAA 133
QY 324 HCFIQHQLMPENVVEKTKAMDHLAVLLEPRPRRPLOMRYSEVSF--MELDKFLEVD 381
DB 134 HAIIVDSKTY--EDVSESAPRIKD-NETIATSKLR---EMAYCQVAAPQVLEKFLDDV 186
QY 382 RNGIYPLMNFPA-----ATRP-----LG 398
DB 187 RNGIYPLTAPALAPLAPQPTTLTAAYSNTNPSTLITTDISAPSSTHPSANSTTAETSVNTS 246
QY 399 LPRVLAAP-----PEEVOK-----412
DB 247 LRGQSSQPSGTTTNTNGPSSIGKSASPEAVDDKKIGVSTESTSKVEVEVNGANVTIGSS 306
```

```

QY 413 ----ATPRPE-----PDSETRKYIQMCNLERSEDKARMLTLLVLEDRLHQ 459
DB 307 NGRDAGSPPEEERBENGERMDRLNRIHLEINVIHLENEE-----MSIVLLEDDQMHRO 360
QY 460 LTYDLLPTSDAODLASELVHYGFLHEDD 487
DB 361 LTTISINKGDNPELTENLTTHGPMQOLD 388

RESULT 10
US-09-862-027-35
; Sequence 35, Application US/09862027
; Patent No. 6558418
; GENERAL INFORMATION:
; APPLICANT: Hodges, Martin R.
; TITLE OF INVENTION: No. 6558418el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 461
; TYPE: PR
; ORGANISM: C. elegans
US-09-862-027-35

Query Match      15.6%; Score 415.5; DB 2; Length 461;
Best Local Similarity 29.9%; Pred. No. 8.9e-32;
Matches 116; Conservative 57; Mismatches 92; Indels 123; Gaps 13;

QY 210 PDLRSPFAERBELNLHFFPEYGEVADGT-AVDIFSGMCALEMV---LEIQNGD 265
DB 14 PDAINHNVKTCRNMKMYHIAPEYENNTLISADIVSFGICSLIAVIGLSGCGNS 73
QY 266 TR--VTEEAIAARSHSLSDPNKREFLLCCLARDPAPRPSAHSLLFRVLFEVHSLKLLAA 323
DB 74 SEGPVTEVDIEKAIKRLLEDPMQODFIRQCLRKDPAPRPSARELLFQIILFEVHSLKLLAA 133
QY 324 HCFIQHQLMPENVVEKTKAMDHLAVLLEPRPRRPLOMRYSEVSF--MELDKFLEVD 381
DB 134 HAIIVDSKTY--EDVSESAPRIKD-NETIATSKLR---EMAYCQVAAPQVLEKFLDDV 186
QY 382 RNGIYPLMNFPA-----ATRP-----LG 398
DB 187 RNGIYPLTAPALAPLAPQPTTLTAAYSNTNPSTLITTDISAPSSTHPSANSTTAETSVNTS 246
QY 399 LPRVLAAP-----PEEVOK-----412
DB 247 LRGQSSQPSGTTTNTNGPSSIGKSASPEAVDDKKIGVSTESTSKVEVEVNGANVTIGSS 306

QY 413 ----ATPRPE-----PDSETRKYIQMCNLERSEDKARMLTLLVLEDRLHQ 459
DB 307 NGRDAGSPPEEERBENGERMDRLNRIHLEINVIHLENEE-----MSIVLLEDDQMHRO 360
QY 460 LTYDLLPTSDAODLASELVHYGFLHEDD 487
DB 361 LTTISINKGDNPELTENLTTHGPMQOLD 388

RESULT 11
US-09-345-473E-34
; Sequence 34, Application US/09345473E
; Patent No. 6558903
; GENERAL INFORMATION:
; APPLICANT: Hodges, Martin
; TITLE OF INVENTION: No. 6558903el Kinases and Uses Thereof
; FILE REFERENCE: 35800/183781
; CURRENT APPLICATION NUMBER: US/09/345,473E
; CURRENT FILING DATE: 1999-06-30
```



```

Db      319  --CFSRIPAGYPIREIIRDCIRVRREERSTVKQL-----LYDDFTPEDLIGIRVEIKN 371
Qy      343  KAMDHLAVLAELPRPRPPLQMKRYSEVSFWELD--KLELVKRGCIYLNMPAATRLGLP 400
Db      372  RDADLNDLNEIOMQLVYDEKQKKQYRFKENEGLOPAPIEN-----414
Qy      401  RVLAPPEEYQAKATPPPEPDSRTVYIOMQCNLSESEDKARHMLTLVLVEDRLHROL 460
Db      415  ----DSDEVVQCMIEQOHIPDETRMITKL-----IKQKV-----DAFRDR 453
Qy      461  TYDLLPTDSADLASELVHYGFLHEDDRMLAA 493
Db      454  DHRLLEIKRAKEEBERIREAEIKELRLABA 486

RESULT 14
US-09-862-027-40
; Sequence 40, Application US/09862027
; Patent No. 6858418
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. 6858418el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 1601
; TYPE: PRF
; ORGANISM: C. elegans
US-09-862-027-40

```

Query Match 13.8%; Score 368.5; DB 2; Length 1601;

Best Local Similarity 25.0%; Pred. No. 2,2e-26; Matches 128; Conservative 88; Mismatches 188; Indels 109; Gaps 20;

```

Qy      10  RAREREREDS-----EDESIDI--LESPPGMRQKREOVNQGMPLOSTPL 57
Db      54  RAKEQEKREBEKKAARRIDVEDDPAQEKIDKSKNGRFLKFDDELGRS---FKTVFR 110
Qy      58  AMDTGEGVYVNNELHFGDKARFAHBEKIQTVEQLVVDHNYVYKHLKTLDTSEACA 117
Db      111  GLDTEGVAVAWCELO--ESKLNKTERQRFREAEMLKLOHNPVIRFYDYW--ESADLCG 167
Qy      118  R--VFITEYVSSGSLKQFLKTKTKK-NHKAMNABAMKRWCTOILSALSFHACSPPIIH 173
Db      168  KRRIYLVLELMTSGTLKMYLKRKINIKV---LKSVCROILKGLSFLHTRNPPVILH 223
Qy      174  GNLTSPTFIQ--HNGLIKIGSVNHRIFSNALRPPTALPDDLRSPIRAREELNLHFFPP 232
Db      224  RDLKCNIFITGTGSKIGDLGLATLKNK-----SPAKSVIGTPPEFMAP 268
Qy      233  EYGEVADGTRAVDFSGCMCLMAVLEI---QTNQDTRVTEEAIAARHSLSDPNKREFI 289
Db      269  EMYEEMVDSVDVYAFGMCLELVNTGEPYSECMNPATYIRKVISGVKPE-----318
Qy      290  LCCLADPARPARSAHSLFPH--RVLPEVNS-LKLLAHCFIQHQLMPENV-----VEEKT 342
Db      319  ---CFSRIPAGYPIREIIRDCIRVRREERSTVKQL-----LYDDFTPEDLIGIRVEIKN 371
Qy      343  KAMDHLAVLAELPRPRPPLQMKRYSEVSFWELD--KLELVKRGCIYLNMPAATRLGLP 400
Db      372  RDADLNDLNEIOMQLVYDEKQKKQYRFKENEGLOPAPIEN-----414
Qy      401  RVLAPPEEYQAKATPPPEPDSRTVYIOMQCNLSESEDKARHMLTLVLVEDRLHROL 460
Db      415  ----DSDEVVQCMIEQOHIPDETRMITKL-----IKQKV-----DAFRDR 453
Qy      461  TYDLLPTDSADLASELVHYGFLHEDDRMLAA 493

```

Db 454 DHRLLEIKRAKEEBERIREAEIKELRLABA 486

RESULT 15

US-09-270-767-31733

; Sequence 31733, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 31733

; LENGTH: 110

; TYPE: PRF

; ORGANISM: Drosophila melanogaster

US-09-270-767-31733

Query Match 13.7%; Score 364.5; DB 2; Length 110;

Best Local Similarity 61.0%; Pred. No. 1e-27;

Matches 72; Conservative 14; Mismatches 21; Indels 11; Gaps 2;

```

Qy      118  RVIFITEYVSSGSLKQFLKTKTKK-NHKAMNABAMKRWCTOILSALSFHACSPPIIHGNT 177
Db      3  RVIFITEYVSSGSLKQFLKTKTKK-NHKAMNABAMKRWCTOILSALSFHACSPPIIHGNT 62
Qy      178  SDTIFIQHNGLIKIGSVNHRIFSNALRPPTALPDDLRSPIRAREELNLHFF--PREY 234
Db      63  CDSIFIQHNGLVKIGSV-----VDDAVHYSVRGRERERERERGAHYFOABEY 110

```

Search completed: January 12, 2006, 11:23:01

Job time : 26 secs

THIS PAGE BLANK (USPTO)


```

Db 209 -----EFIOKCLQSEPARPTARELLFHPALFEVP 238
Qy 317 SLKLLAHCIOHOYLMPEVNEVEKTKAMDLHAVLAELRP-PRRPLQOKRYSVSMELD 375
Db 239 SLKLLAHCIOHOYLMPEVNEVEKTKAMDLHAVLAELRP-PRRPLQOKRYSVSMELD 286
Qy 376 KFLVDVNGIYPLMNAATRLGLPRVLAAPPEEV-----QKAKTPTEPPDSETRKYI 429
Db 287 -----TLNGIYPLTA-----GLPRPQOQOEVSVPVPSVKITPEPRAVEITRYKV 335
Qy 430 OMOCNLERSEDKARMLTLTLVIEDRLHROTLYDLLPTBSAODLASELYHGFLEH 485
Db 336 LMOCNIESVEEGVGHHTLTLKLEDKLNRHSLCLMPNENIPELAELVQLGFISE 391

```

RESULT 2

```

US-60-752-355-38255
; Sequence 38255, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 38255
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Oryza sativa
US-60-752-355-38255

```

Query Match 7.5%; Score 199; DB 8; Length 554;

Best Local Similarity 22.8%; Pred. No. 1.1e-10; Matches 115; Conservative 66; Mismatches 167; Indels 154; Gaps 25;

```

Qy 2 AAEPAAPRAREREREREDESEDESILLESPPC-----GRWQKR-----EQVNOG 47
Db 125 SSSHPPLRP-----PSSPSSSRVYSSQKKKXKLGSFGVYQO 165
Qy 48 -NMPGLQSTFLANDTEBGEVYVNNELHFGDKAPAAHEEKIQTVFQOLV---VDHPNIV 103
Db 166 FNSGGQ-----MCAIKVAVISDD-----SNSKECLROLHOEIVLSQLSHPNIV 211
Qy 104 KLMHWLDTSEACARVIFITEYVSSGSLKQFLKTKKNHKAAMARAKMCTQILSALP 163
Db 212 QY--YSGDLSSTLSYTL--EYVSGSILKLOE---YAFGEAVLRNTTAQILSGLAV 263
Qy 164 LHAESPPIHGNLTSDTIFIOHNGLIKI---GSVWHRIFSNALRPPTALPDDLRSPIRAE 220
Db 264 LHGRN--TVHRDIKGANILVDPNGDIKLADFGMAKH-----ISAHTSIKSPKGSYYMA 315
Qy 221 REELRLNHFFRPPEYGEVADSTAVDIFSFGCALEMAVLE---IQNGDTRVTEBAILARAR 277
Db 316 PEVIMNTNGY-----SLSYDWSLGGTIEMTARBPWIOYSG-----VAAIFKIG 361
Qy 278 HSLSDPNM-----REFILCCIARDPARPSASHILFRVLFVY----- 315
Db 362 NSKOIPDIPDHLSEAKNPLKLCQORDPARPPTAOLMEHPYKDYLVANRSPSGMTRDT 421
Qy 316 -----HSLKLLAHCIOHOYLMPEVNV-----EEKTKAMDLHAVLAELRP 355
Db 422 FPTSPOGKAMVQTSNRSLSPRLDPDLTMNLQVPTSAISISTRISANPNVMMNS 481
Qy 356 RPRRP---PL-QNRYSEVSFMELDKF-----LEDVNGIYPLMNAATRLGLPRV 402
Db 482 LPVSPGSSPLRQYRQSNRSCMPSPPHPAYSAGANYSINNALYP-----KRP---SNH 532
Qy 403 LARPEEVOKAKTPTPEPPDSETR 426
Db 533 VDPWLEISQLKAP---PFDSPRR 553

```

RESULT 3

```

US-60-752-355-5150
; Sequence 5150, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 5150
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (695)..(695)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; NAME/KEY: misc.feature
; LOCATION: (726)..(726)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
US-60-752-355-5150

```

Query Match 7.3%; Score 196; DB 8; Length 833;

Best Local Similarity 25.3%; Pred. No. 4e-10; Matches 83; Conservative 61; Mismatches 112; Indels 72; Gaps 17;

```

Qy 32 SPGGRWQKRREGVNOGNMFGLOSTF---LAMDTEBGEVYVNNELHFGDKAPAAHEE 85
Db 404 SPGSRWK-----GKLIG-RGTFGHYVVGNSDGEWCAKKEVTLFDDPKS-----XE 451
Qy 86 KIQTVFQOLV---VDHPNIVLHKWMDTSSACARVIFI--TEYVSSGSLKQFLKTKKN 141
Db 452 SAKOLGOEISLSRLQHNPVIQYVY-----GSETVDKLYILEYVSGSILKLOE--- 502
Qy 142 HKAMNARAKMCTQILSALPILHACSPPIHGNLTSDTIFIOHNGLIKIGSVWHRIFSN 201
Db 503 YQGLGEQALRSTYQOILSLAYLHAKN--TVHRDIKGANILVDPGSRVTLADFGMAKHIN 560
Qy 202 ALRPTALPDDLRSPIRAEELRLNHFFRPPEYGEVADG--TAVDIFSFGCALEMAVLE 259
Db 561 GQCCPSPFKG--SP-----YMAPEVINKNSNCNLAVDIWSLGGTYLEWATSK 606
Qy 260 ---IQNGDTRV---TEBAILARHSLSDPMREBILCCIARDPARPSASHILFRVLY 312
Db 607 PMSOYEGIAAMFKIGNSKELRPIPDHLEPG-KOFTIRKCLORDPSQRPAMELL----- 660
Qy 313 FEVHSLKLLAHCIOHOYLMPEVNV 340
Db 661 -----QHPVQKAVSLKSVLSE 678

```

RESULT 4

```

US-60-752-355-13134
; Sequence 13134, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 13134
; LENGTH: 894
; TYPE: PRT
; ORGANISM: Oryza sativa (japonica cultivar-group)
US-60-752-355-13134

```

Query Match 7.3%; Score 196; DB 8; Length 894;

Best Local Similarity 25.3%; Pred. No. 4.5e-10;

US-60-752-355-1875

Query Match 7.3%; Score 195.5; DB 8; Length 611;

Best Local Similarity 25.0%; Pred. No. 2.8e-10;

Matches 108; Conservative 57; Mismatches 134; Indels 133; Gaps 24;

QY 36 RMOKREOVNOGMPGLQSTF---LAMDTEEG-----VEVNMELHFGDKAPAAIE 84

DB 206 KMKK-----GKLLG-RGTFGVVYLGFRNQGOMAKKEVRVSD-----QTKK 248

QY 85 EKIQTVEQDLV---VDHPNIVKHLKXWLDTSEACARVIFTEYVSSGLKQFLKTKTKN 141

DB 249 ECKQJNQEIILSNLSHPNIVVYVGSLEDELTVL---EYVSGSIIHKLQE----- 300

QY 142 HKAMBARAKMCTQTLNLSLSEAFHACSPRIHGNLSDTFIOHNLIKI---GSVWHRI 198

DB 301 YGAFREPVQNYRQLHLSLGFHANN--TVHNDIGANLVDNPBEIKLADFGMAKH-I 357

QY 199 FSNALRPPTLPDDLSPIRAREEDLNHFFPPEYGEVADGTVADIFSGMCALEMA-- 256

DB 358 TSSSL-----VLSFKSPYMAAEVVMNTSGY-----GLPVIWISLGCALILEMASS 403

QY 257 -----VLEIQNGDTRKVTBEAIPARAHSLSDPMMEFLICCLADPARPSASHLL 307

DB 404 KPFWSQYEGVAAIFKIGNSKDFPEI---PDHLSND--AKKFILCLQREPSARPTASQLL 458

QY 308 FHRVLFVHSLKLLAHCFIQHOYLMPENVVEKTKAMDLHVLALPRP---RRPPIQ 363

DB 459 -----EHPEVKNQ-----STTVTHGVTKAYPPSPDGNRTPEV- 493

QY 364 WRXSEVSMELDKLELDVNGIYPLM-NFAA-----TPPLGPRVLAAPPEVEQAKTP 416

DB 494 -----LDSGGRNISPTKGNVASHPIIISRLSCPR-----EIVTTIS 532

QY 417 TP-EPPDSETRK 427

DB 533 LPVSPITSSPLRQ 544

RESULT 8

US-60-752-355-42688

Sequence 42688, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: Patentin version 3.3

SEQ ID NO 42688

LENGTH: 567

TYPE: PRT

ORGANISM: Oryza sativa

US-60-752-355-42688

Query Match 7.2%; Score 193.5; DB 8; Length 567;

Best Local Similarity 25.3%; Pred. No. 3.9e-10;

Matches 86; Conservative 57; Mismatches 107; Indels 95; Gaps 17;

QY 2 AABEPAPRAREREREREDESESDILESPPC-----GRMOKER-----EQVNOG 47

DB 203 SSGHPLRP-----PGSPCSSSRVSSQMKKGLLOSCTFGQVYQG 243

QY 48 -NMPGLQSTFLANDTEGVEVNMELHFGDKAPAAHEKIQTVFQDLV---VDHPNIV 103

DB 244 FNSGCG-----MCAIKKAVISDD-----SNSKECLRLQHQEIVLLSQLSHPNIV 289

QY 104 KHLKWLDTSEACARVIFTEYVSSGLKQFLKTKTKNKKAMARAKMCTQTLNLSLS 163

DB 290 QY--YSDLSSETLSVYL--EYVSGSIIHKLQE---YAGFBAVLRYNTAQIISGLAV 341

QY 164 LHACSPPIIHGNLSDTFIOHNLIKI---GSVWHRIFSNALRPPTALPDDLRSPIRAE 220

DB 342 LHGRN--TVHNDIGANLVDNPBEIKLADFGMAKH-----ISATSIKSPKSGPYMA 393

QY 221 REELNHLHFPPEYGEVADGTVADIFSGMCALEMAVLE---IQNGDTRVTEAIPAR 277

DB 394 PEVIMNTNGY-----SLSDIWSLGCCTIIEAATRAPPIIOYEG-----VAAIFKIG 439

QY 278 HSLDDPM-----REFILCLADPARPSASHLLFH 309

DB 440 NSKIDPDIPLHLSFAKNFLKCLQRPAPARTAAQLMKH 479

RESULT 9

US-60-752-355-35913

Sequence 35913, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: Patentin version 3.3

SEQ ID NO 35913

LENGTH: 1484

TYPE: PRT

ORGANISM: Candida albicans SC5314

US-60-752-355-35913

Query Match 7.2%; Score 193; DB 8; Length 1484;

Best Local Similarity 22.6%; Pred. No. 1.9e-09;

Matches 79; Conservative 56; Mismatches 154; Indels 60; Gaps 13;

QY 10 RAREEREREREDESESDILESPPC-----RMOKREOVNOGMPGLQ 53

DB 1135 KIEERRRNLQEBFQSIGRLVDSDEYQFLTLASSFSSVSIRWQ-GACIGRT---FG 1190

QY 54 STEFLAMDTGEGVAVNMELHFGDKAPAAHEKIQTVFQDLVHNPNI-----VTLHKY 108

DB 1191 QVFSVNLDTGGVMAVKETTFHDQSQVKTIVSIEEMKTVLMLNHPNVQYFGVEVHR- 1249

QY 109 WLDTEACARVIFTEYVSSGLKQFLKTKTKNKKAMARAKMCTQTLNLSLFLHACS 168

DB 1250 -----DKVYIFMEFCGSLAGL---THGRJEDEMVIOYVALQWLBGLAYLHQSG 1297

QY 169 PRILGNLSDTFIOHNLIKI---IGSVWHRIFSNALRPPTALPDDLRSPIRAEBEIR 225

DB 1298 --VHRDIPENVLLDHNGVIXYVDFGA-KVIASGRITIGMTSSLSKSVKRDGNML 1354

QY 226 NLHFPPEY--GEVADGTA-----VDIPSGMCALEMAVLE---IQNGDTRVTEAIA 274

DB 1355 NSMTGTPMTSPBAITGISTDSGVVDIWSLGCCTIIEAATRAPPIIOYEG----- 1414

QY 275 RARHSLSDPM-----REFILCLADPARPSASHLLFHRVLFVHSL 318

DB 1415 GHKPOLPSPDEQSEPORTFLARCLEHDPKRSASAVELADPMWVEIRHM 1463

RESULT 10

US-60-752-355-24004

Sequence 24004, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: Patentin version 3.3

SEQ ID NO 24004

LENGTH: 786

TYPE: PRT

ORGANISM: Zea mays

US-60-752-355-24004

Query Match 7.2%: Score 191.5; DB 8; Length 786;

Best Local Similarity 21.5%; Pred. No. 1e-09; Indels 161; Gaps 26;

Matches 109; Conservative 78; Mismatches 160; Indels 161; Gaps 26;

QY 2 AAEPAARRAREBEREREDESESDILEESPCC---GRWQKREOVGNMPC---LQ 53

DB 356 SSGQLP-----LPGSPCLPSSSLQMK-----GKLGSSTPG 369

QY 54 STLAMDTEG-----VEVWNEHLFGGRKAPAAHEEKIQVFEQVLV---VDHENV 103

DB 390 QVWGFSEGGOMCAIKKVVISD-----SNKESLRQLNGRIVLLSQSHPNIV 440

QY 104 KLRKYMWDTEACARVFI--TEYVSSGSLKQFLKTKKKNKAMARAKWCTQIISLS 162

DB 441 QYV-----GSDLCNETSVYLEYVSGSIHKLQOE---YGPGEAVLRYVTAQIISGLA 491

QY 163 FLHACSPPIHGNLTSPTFIQHNLGIKI---GSVWHRIFSNALRPPTALPDRLSPIRA 219

DB 492 YLHGRN--TVHRDIKGNILVDPNGDIKLADFGMAK-----ISATYISKSGSPYMM 543

QY 220 EREELRNHLFPPEYGEVADGTAVDIFSPGMCALMAVLE---IQTNGDTRVTEBAYARA 276

DB 544 APFVIMNSNGY-----SLSDIWSLCTILEMATAKPPMSQYEGVAIFK--IGNS 592

QY 277 RHLSDPN---MREFIICCLARDPARRSASHLSLFH----- 309

DB 593 KQIPDIPNLSSEAKGFLKLCLODPAPAPTAQMLDHPVKQATVRSRSSITRDMFP 652

QY 310 -----RYLFEVHSLKLAACHFIQHQLAMPENV-----VEEKTAMDHLAVL 351

DB 653 NSTDGKNSRKVIETSSRSLSP---LRDPILGRNLPGPSPISPSRRIALNLSNVR 709

QY 352 AELPRRRP---PL-QMRSEVSFMEIDKF-----LEDVRNGIYPLMNAATRPPLG 398

DB 710 MNMSLVSPSSPLRQYRQGNRSCLSPRPAYSAGANYPINNALYF-----TRP-- 761

QY 399 LPRVLAPPEEVOKATPTPEPDSER 426

DB 762 -SSGLTDPWLIEISQVKTQT---FDSPPR 785

RESULT 11

US-60-752-355-10912

Sequence 10912, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: PatentIn version 3.3

SEQ ID NO 10912

LENGTH: 627

TYPE: PRT

ORGANISM: Lotus corniculatus var. japonicus

US-60-752-355-10912

Query Match 7.1%: Score 190.5; DB 8; Length 627;

Best Local Similarity 24.9%; Pred. No. 8.8e-10; Indels 101; Gaps 21;

Matches 106; Conservative 57; Mismatches 162; Indels 101; Gaps 21;

QY 28 ILEESPC--GRWQKREOVGNMPCGLQSTF---LAMDTEGVEVWNEHL--FGDRKAF 80

DB 210 VLENNCTNLSKMKK-----GKLGG-RGTFGHVYLGFPNENGMCAIKKVFSDK-- 259

QY 81 AAHEEKIQVFEQVLV---HPNIVKLAHKYMDTSEACARVFTIEYVSSGSLKQFLPK 137

DB 260 -TSKECLKOLNOLNINLNFSPHIVQY--YGESELEESLSVYL--EYVSGSIIHKLQOE 314

QY 138 TKNHKAARAKWCTQIISLSFLHACSPPIHGNLTSPTFIQHNLGIKISVWHR 197

DB 315 -----YGAFFEPVQNTROQVSGLAYIH--SRATYVRDIKGNILVDPNGEITLADPG-- 366

QY 198 IFSNALRPPTALPDRLSPIRAREBEURLNHLFPPEYGEVADGTAVDIFSGMCALEMA- 256

DB 367 -MSKHINSAASMLSFKSGPYWMAPEVWNTNGY-----GLPVDISLGICTIEMAT 416

QY 257 -----VLEIQNGDTRVTEBAYARHSLSPNMRKFIICCLARDPARPSHSL 306

DB 417 SKPPWQFEGVAALFKIGNSKMP---LPEHLSD--DAKNFTKQCLQRPDLARPTAQSL 471

QY 307 LFRVLFEVHSLKLAACHFIQHQLAMPENVEEKTAMDHLAVLEP---RPRRPL 362

DB 472 L-----NHPTRDQ-----SARKVANASTTRDAFPMSGSRTPV 507

QY 363 QMRSEVSFMEIDKFLEDVRNGIYPLMNAATRP--LGLPRVLAPPEEVOKATPTPEP 421

DB 508 LEPHSN-----RSSITLDVATKPALAAVTLNPRDSTRTITSLPSPS 554

QY 422 DSETRK 427

DB 555 SSELRLQ 560

RESULT 12

US-60-752-355-30838

Sequence 30838, Application US/60752355

GENERAL INFORMATION:

APPLICANT: Abad, Mark et al.

TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

FILE REFERENCE: 38-21(53720)B

CURRENT APPLICATION NUMBER: US/60/752,355

CURRENT FILING DATE: 2005-12-21

NUMBER OF SEQ ID NOS: 52803

SOFTWARE: PatentIn version 3.3

SEQ ID NO 30838

LENGTH: 555

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-60-752-355-30838

Query Match 7.0%: Score 186.5; DB 8; Length 555;

Best Local Similarity 23.3%; Pred. No. 1.8e-09; Indels 101; Gaps 20;

Matches 97; Conservative 55; Mismatches 164; Indels 101; Gaps 20;

QY 42 EQVGNMPCGLQSTFLAMDTEGVEVWNEHLF---GDRKAPAAHEEKIQVFEQVLVD 98

DB 8 EQGKGS-----FSGALLVRHKQEKRYVLKKIRLARGSDARSAHQE-----MELISTVR 59

QY 99 HPNIVKLAHKYMDTSEACARVFTIEYVSSGSLKQFLKTKKKNKAMARAKWCTQI 158

DB 60 NPEVVEYKDSWVE--KGC--YVCIVIGYCOGDMTDTIKRACGVH--FPEEKLQWMLVQL 114

QY 159 SALSLFLHACSPPIHGNLTSPTFIQHNLGIKISVWHLIFSNALRPPTALPDRLSPR 218

DB 115 MALDYH--SNHILHADVCSNIFLTKEDIRLGD-----FGLA--KITLSODLTSSVV 164

QY 219 AEREELRNHLFPPEYGEVADGTAVDIFSGMCALEMA-----VLEIQNGDTRVTEB 271

DB 165 GTPS-----YMCPELLADIPIYGSKSDIWSLGCMTMAAHKPPFKASDVQT----- 210

QY 272 AIAARASHLSDP-----MREFIICCLARDPARRSASHLSLHFRVLEVHSLKLAH 324

DB 211 LITTKIHKLIMDPIPAWYSGSFRGLIKSMRLKPELRPSANEHL----- 253

QY 325 CFIOHQLAMPENVEEKTAMDHLAVLELPRRRP--PLQMRYSVSFMEIDKFLEDVR 382

DB 254 ---NHPLOPY-----ISWYMKLESBSPRSTFLO-----PSEBDATLKERR 292

QY 383 NGIYPLMNAATRPPLGLPRVLAPPEEVOKATPTPEPDSERTRKVTQMOCNLESE 439

DB 293 RS-----SFSNDRRLN--PVSVDTEAGSVSSGKASPTPW--FNGRKVSEYTVGVREE 342

Query Match 7.0%; Score 186; DB 8; Length 544;
 Best Local Similarly 24.3%; Pred. No. 1.9e-09;
 Matches 81; Conservative 57; Mismatches 115; Indels 80; Gaps 15;

```

RESULT 14
US-60-752-355-51629
; Sequence 51629, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 51629
; LENGTH: 884
; TYPE: PRT
;

```

Query Match	7.0%;	Score 186;	DB 8;	Length 884;
Best Local Similarity	24.7%;	Pred. NO. 4e-09;		
Matches	84;	Conservative	60;	Mismatches 116;
				Indels 80;
				Gaps 17

Db 653 -----DHAFVRNVMERPIVSGBPAEAMNV 678

US-60-752-355-18183

Matches 84; Conservative 60; Mismatches 116; Indels 79; Gaps 17

Db 553 QSGPLSFKG---SP-----YWMAPEVIKNSNGSLAVDIWSLGLTVLEMATTKP 598

Qy 260 --IQTNGDTRVTEATARARSLSDPM-----REFILCCLARDPARPSAHSILFH 309
 Db 599 PMSQYEG----VPAMFKIGNSKELPDIPDHLSEBEGDFVRKCLQRNPNRPTAQL-- 651
 Qy 310 RVLFEVHSLKLAHCFIOHQYLMPEVVE-EKTKAMD 347
 Db 652 -----DHAfVRNVMPMERPIVSGEPANMNV 677

Search completed: January 12, 2006, 11:30:50
 Job time : 50 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using bw model

Run on: January 12, 2006, 11:17:30 ; Search time 18 Seconds
(without alignments)
2710.105 Million cell updates/sec

Title: US-10-618-941-99

Perfect score: 2670
Sequence: 1 MAAPBPAPRRARRRRERED.....RMKLAFLBSTFLKRGCTQA 507

Scoring table: BL0SUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80:*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1346	50.4	258	2 T46491	hypothetical prote
2	543	20.3	231	2 T23136	hypothetical prote
3	415.5	15.6	461	2 T23574	hypothetical prote
4	372	13.9	516	2 T06716	hypothetical prote
5	368.5	13.8	1851	2 T19964	hypothetical prote
6	349.5	13.1	677	2 T02951	hypothetical prote
7	327.5	12.3	547	2 T46059	probable mitogen a
8	327.5	12.3	549	2 T51099	MAP kinase (import
9	316.5	11.9	633	1 S49611	probable serine/th
10	195	7.3	1401	2 T39225	MAP kinase kinase
11	191	7.2	544	2 S40482	serine/threonine-s
12	190	7.1	366	2 T51339	mitogen-activated
13	190	7.1	345	2 G01773	p21-activated prot
14	188	7.0	525	2 S58682	protein kinase, p2
15	187.5	7.0	544	2 A57597	beta-p21-activated
16	186	7.0	378	2 T26684	hypothetical prote
17	185.5	6.9	544	2 T49376	p21 activated kina
18	185.5	6.9	883	2 A96662	hypothetical prote
19	185	6.9	738	2 F96701	hypothetical prote
20	183	6.9	1895	2 T06609	disease resistance
21	181.5	6.8	553	2 T01479	hypothetical prote
22	181	6.8	480	2 A86371	hypothetical prote
23	180.5	6.8	963	2 T09911	probable serine/th
24	177.5	6.6	342	2 T04054	protein kinase hom
25	177.5	6.6	736	2 T05137	protein kinase hom
26	176.5	6.6	403	2 J05974	autora-related kin
27	175	6.5	481	2 T49072	protein kinase - m
28	174.5	6.5	608	2 G96575	probable MEK kinas
29	171	6.4	1008	2 D84434	probable receptor-

30	170	6.4	372	2 T52621	mitogen-activated
31	170	6.4	391	2 T48115	protein kinase ATM
32	170	6.4	528	2 D86456	probable protein k
33	167.5	6.3	348	2 T51340	mitogen-activated
34	166.5	6.2	348	2 T52635	mitogen-activated
35	166.5	6.2	560	2 D85084	probable mitogen-a
36	166.5	6.2	746	2 S62365	SNF1-related prote
37	166	6.2	338	2 J07838	Nek6 protein kinas
38	164.5	6.2	487	2 A71407	probable Ste20-lik
39	163	6.1	1895	2 T15881	hypothetical prote
40	162.5	6.1	462	2 S29851	protein kinase 6 (
41	162.5	6.1	535	2 T51736	mitogen-activated
42	161.5	6.0	406	2 T52626	probable mitogen-a
43	161.5	6.0	706	2 A48084	SRB1 protein kina
44	161	6.0	722	2 T37970	probable G2-specif
45	160.5	6.0	677	2 T27127	hypothetical prote

ALIGNMENTS

RESULT 1
T46491
hypothetical protein DKFZp434P086.1 - human
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46491
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
Submitted to the Protein Sequence Database, January 2000
A:Reference number: 223037
A:Accession: T46491
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-258 <AAA>
A:Cross-references: UNIPROT:Q9NSY0; UNIPARC:UP1000006E116; EMBL:AL137662
A:Experimental source: adult testis; clone DKFZp434P086
C:Genetics:
A>Note: DKFZp434P086.1

Query Match 50.4%; Score 1346; DB 2; Length 258;
Best Local Similarity 100.0%; Pred. No. 2.6e-81;
Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	250	MCALENAVLEIQNGDTRVTEBAIARARHSLSPNNREFTLLCTIARDPAPRPSAHSLLFH	309
DB	1	MCALENAVLEIQNGDTRVTEBAIARARHSLSPNNREFTLLCTIARDPAPRPSAHSLLFH	60
QY	310	RVLPEVHSLKTLAAHCFIOHQYLPENVVEKTKAMDHLAVLAELPRRRPPIQWRYSEV	369
DB	61	RVLPEVHSLKTLAAHCFIOHQYLPENVVEKTKAMDHLAVLAELPRRRPPIQWRYSEV	120
QY	370	SPFELDKFLDVANGIYPLNPNFATRPGLPRVLAPPBEVQAKTPTPPPSSETRKVI	429
DB	121	SPFELDKFLDVANGIYPLNPNFATRPGLPRVLAPPBEVQAKTPTPPPSSETRKVI	180
QY	430	OMOCNTERSDKARWHITLLVLVBDRILHROLTYDLPDTSQAQDIASELVHYGLHEDDM	489
DB	181	OMOCNTERSDKARWHITLLVLVBDRILHROLTYDLPDTSQAQDIASELVHYGLHEDDM	240
QY	490	KLAAPLESTFLKRGCTQA 507	
DB	241	KLAAPLESTFLKRGCTQA 258	

RESULT 2
T23136
hypothetical protein H37N21.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23136; T23575
R:McMurray, A.
Submitted to the EMBL Data Library, November 1997
A:Reference number: Z19694

QY 17 ERDESEDSILLESPCGRWQKRQEVQNOGMFGLOSTFLAMTBEQVEVYNNMELHFGD 76
Db 2 EISSASDSISIAVETDPDSGRYGRFEVEVIGKAM--KTVYAFQVQVLGMEVANNQVVKNE 58
QY 77 RKAFANAHEKIQTVESQVL--VDHPNIVKLHKYMLDPTSEACARVIFITEVYSSGLKQ 133
Db 59 --VFRS-PEPIQRLYSEVHLKLNANHESTIRYCTSMIDVNRRTN--FITLEFTSGTLRE 113
QY 134 FLKTKTKNKKAMNANAMRWCTOILSALSFHACSPPRIHGNLTSDTIFIO-HNGLIKIG 192
Db 114 YRRKYQK----VDIRAKISWARQILNGLAYLHGHPPIVHRDLKCDNI FVNGHILGYKIG 169
QY 193 SWMHRIFSNALRPPALPDILRSPTRAREELRNH-----FPPRYGEVADGTAVDI 245
Db 170 DL-----GLAALIRGS-----QNAHSVIGTPEFMADELVEEDYNELVDI 208
QY 246 FSPFGCALEMAVLEIQTNQDGT-----RYVEEALIRARHSLSPNNKREFLCTCLADPA 298
Db 209 YSPGCVLEMLTGEYRSECTNPAQIYKRVISGKLPSFHLIQHTEAQRFYGKGL-BTVS 267
QY 299 RPSASHSLFPRVLPEVHSLKLAAHCFIQHQYLMPENVSEKTKAMDHLAVLALELPRR 358
Db 268 RRLPAKELLADPFLAATDERDL--APLFRLLQOLAIOQLANGT-----VGHLPSTT 318
QY 359 RPPLOWRYSEVS-----FMELDKLELD--VANGIYPLMNEFAATPGLCPRLVA 404
Db 319 DPT---RTDMSINGKMSSEDTTIFLQVILDGSGHMRNIQFPF-----NILS 363
QY 405 PPPEVQCAKTPTPPEPFDSETRKVIQWQCNLERSEDKARWH-LTLLVLVEDRLHQLTYD 463
Db 364 DTPLELV-----ALEWKELEITD-----MDPLEIAMIEHISL----- 397
QY 464 LLPTDSADLASSELVHYGLHEDD 487
Db 398 LVPWNRAND--SSIRHESFGHEDD 419

RESULT 8
T51099
MAP kinase [imported] - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C|Accession: T51099
R|Cooke, R.M.; Laudie, M.; Berger, C.; Delserny, M.
submitted to the EMBL Data Library, May 1999
A|Description: Analysis of Arabidopsis thaliana gene structure by cognate cDNA sequencing
A|Reference number: Z25297
A|Accession: T51099
A|Status: Preliminary; translated from GB/EMBL/DBJ
A|Molecule type: mRNA
A|Residues: 1-549 <COO>
A|Cross-references: UNIPROT:O9XFS5, UNIPARC:UP100000A62F3, EMBL:AJ238802, PIDD:CA843520

Query Match	12.3%	Score 327.5	DB 2	Length 549
Best Local Similarity	25.8%	Pred. No. 4.9e-14		
Matches 130	Conservative 79	Mismatches 176	Indels 119	Gaps 23
Qy	17 EREDESEBESDILLESPOCRWQKRREQVQCNMPGOLSTFLAMTEGEGEVVNNELHFGD	76		
Db	4 EISSSDSIAYVEIDPESGRYGRFREVLDGKAM---KTVYKAPQVYLGEMVANNQVATNE	60		
Qy	77 RKAPAEERKIQVFEQLV---VDHENIVKLHKMYLDTSACARVIFITEYVSSGLKQ	133		
Db	61 --VFRS--DEPQRLRSEVHLKKNLHSEIIKYCTSMIDVNNRTFN--FITELEFSTGLRE	115		
Qy	134 FLKKTKKHKMANNAARWKCQTQILSALSFACSPPIHGNLTSDIFIO--HGLKIKIG	192		
Db	116 YRKTKQK---VDIRAKTSMAKQILINGLAVYHGHDPVIRHDKCDNIFVNGHIGQVYKIG	171		
Qy	193 SVWHRIFSNALRPPTALPDCLRSPRIAREEELRNLIH-----FPPYEGVADGTAVTI	245		
Db	172 DL-----GLAALIRGS---QNAHSVIGPPEMAELVEEDNELVDI	210		

[illegible]

RESULT 9

S49611

probable serine/threonine-specific protein kinase pkpa (EC 2.7.1.-) - Phycomyces blakeslii

C|Species: Phycomyces blakesleeanus

C|Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C|Accession: S49611; S59578

R|Ruiz-Perez, V.; Murillo, F.; Torres-Martinez, S.

submitted to the EMBL Data Library, November 1994

A|Description: A novel serine/threonine protein kinase in the fungus Phycomyces blakeslii

A|Reference number: S49611

A|Accession: S49611

A|Molecule type: DNA

A|Residues: 1-633 <RUI>

A|Cross-references: UNIPARC:UPI0000172621; EMBL:Z46636

R|Ruiz-Perez, V.L.; Murillo, F.J.; Torres-Martinez, S.

Curr. Genet. 28, 309-316, 1995

A|Title: PKPA, a novel Phycomyces blakesleeanus serine/threonine protein kinase.

A|Reference number: S59578; MUID:96120859; PMID:8590476

A|Accession: S59578

A|Molecule type: DNA

A|Residues: 1-9;38-346 <RUM>

A|Cross-references: UNIPARC:UPI0000172622; UNIPARC:UPI0000172623; EMBL:Z46636

A|Note: only part of the coding region is given

C|Genetics:

A|Gene: pkpa

A|introns: 60/1; 248/3; 277/2; 339/3; 534/1

C|Superfamily: Phycomyces blakesleeanus probable serine/threonine-specific protein kinase

C|Keywords: phosphotransferase; protein kinase

F|34-289|Domain: protein kinase homology <KIN>

Query Match 11.9%; Score 316.5; DB 1; Length 633;

Best Local Similarity 23.7%; Pred. No. 3.1e-13; Indels 125; Gaps 23;

Matches 126; Conservative 87; Mismatches 193;

15 EREREDSE--DESDILESPCGMQRKREQVANGNPFGLQSTFLAMDTEEGEVVNNEL 72

12 DMEVVDSDSNMPTVEKVIENASGNRGYSKLTNVLDKG---AKVVKYKIDREAI----NDN 64

73 HFGDRKFAAHEEKIQTFEQLVLDHPNIVKLHKYKLTDTSEACARVIFTEYVSGSLK 132

65 EITVVKTRQBFKGLGHEIDILKSVRHPNITTFHDAAVNNETE---FVFITELMTSGTLR 120

133 QFLKTKGNKHAMNARMKRCQIILSALSFLLACSPRIHGNULTSPITIQ-HNGILIKI 191

121 EYIRKLT--PLPIKIKVKKMCRQILKGLAVLHGHEPPIIHRDIKCNINIFNGAHGEIKI 177

192 GSVNHRIFSNALRPPTALPDILRSPFARREELRNLIFFPEE-YGEVAQDSTAVIDFSFGM 250

178 GDMGTALMKNGKQKTYV-----IGTP-----EFMAPMYEEGYNKVDIYAFGM 221

251 CAEMAVLEIQTNGDT-----RVTEEARARAHSLSDNMKEFFIICLARDPARRPSA 303

Db 222 CLEMATGEYGESECTNANVQFKKVTQTIKRECLSRVQDELLTVNICTLPED-ERM7A 280
 304 HSLFHRVLE---EV-----HSLKTLAANCF-----IQHY-----LM 333
 Db 281 QETLEHREFLAVEREVLVSKDMTKLTLQVFKMDKLSVKTEFNADTDADVAEMI 340
 334 PENVVE---BKTAMDIAVLAEIPRRPP-----LQMRSEVSFMELDKFLEEDVRN 383
 Db 341 EEOVLQNCYQQLTCEINRLRDARNOQSPDKGDEKTYWRENDRSLEKAKDIAL 400
 384 GIVPLAN-----FAATR-PLGLPRVLAPEEVOKATPTPE 419
 Db 401 AVERFEAEKCELLBQNIIEBCKETIFALEQAKFGQPDLLQPPQ---QOPQOPQ 457
 420 PPSERTRKVIQMCNLE-----RSEDKARHLLVLLVLERLHR 458
 Db 458 P---QOPQOPQLOPOLQVLSPGSTSPGPTSDNST-NSTMLSLSELSK 504

RESULT 10

T39225
 MAP kinase kinase kinase - fission yeast (Schizosaccharomyces pombe)
 C/Species: Schizosaccharomyces pombe
 C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-Oct-2004
 C/Accession: T39225
 R/Churher, C.M.; Gentiles, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 Submitted to the EMBL Data Library, August 1997
 A/Reference number: Z21837
 A/Accession: T39225
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: DNA
 A/Residues: 1-1401 <CHU>
 A/Cross-references: UNIPROT:O14299; UNIPARC:UPI0000138F13; EMBL:Z98763; PIDD:CA811500.1;
 A/Experimental source: strain 972h-; cosmid c961
 A/Genetic:
 A/Map position: 1

Query Match 7.3%; Score 195; DB 2; Length 1401;
 Best Local Similarity 25.1%; Pred. No. 8.1e-05;
 Matches 87; Conservative 50; Mismatches 135; Indels 74; Gaps 18;

36 RMQKRRQVQGMN--PGL-QSTFLAMDTEGVEVVNNEHFGDRKAFANHEKIQVFE 92
 Db 1036 RWO-----QGHFVRSQMGFDVYTVNMTGDLAVKEIKLQDSRFRSTVQJHNEMT 1088
 93 QLVLVDPHPI-----VLAKYVLDTSBACARVIFITEVSSGLKQFLKTKKXHKAMNA 147
 Db 1089 VLRLNHPNVVTVYGVEVHR-----EKVYTFMEFCGGSLADL---AHGRIDE 1135
 148 RANKRWCTQILSALSFHACSPPIHGNLTSDTIFIQHGLIKIGVWHRIFSNAL--RP 205
 Db 1136 NVLKVVYVQLLEGLAYIH--SQHILHRDIPANILDRGMKIKSD-----FSGALYVSP 1188
 206 PTLPPDLLRSPIRAREELNHPPEY--GEVADT-----AVTIFSGCALEMAV 257
 Db 1189 PTL-PEVRYEDIOPELOHLAG---TPMYAPETIILTKKGDPCAMDIWSLGVILEM-- 1241
 258 LELQNGDTRVTEAL---ARARHSLSDPN-----MREPIICCLARDPARPSAHSIL 307
 Db 1242 MTGSTPSEMDNEKALVYVAAMHTPSIPQNEKISLARDPIECGFRDEQRPRAVDL 1301
 308 FHRVLFVHSLKTLAAHCFIOHOYLMBENVVEKTKAMDLAAVLAE 353
 Db 1302 THPWITDFRKKTIT-----MPATITKTSLS--HTIRE 1335

RESULT 11

S40482
 Serine/threonine-specific protein kinase (EC 2.7.1.-) - rat
 C/Species: Rattus norvegicus (Norway rat)
 C/Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 05-Oct-2004
 C/Accession: S40482

R/Manser, E.; Leung, T.; Salihuddin, H.; Zhao, Z.; Lim, L.
 Nature 367, 40-46, 1994
 A/Title: A brain serine/threonine protein kinase activated by Cdc42 and Rac1.
 A/Reference number: S40482; MUID:94150586; PMID:8107774
 A/Accession: S40482
 A/Molecule type: mRNA
 A/Residues: 1-544 <MAN>
 A/Cross-references: UNIPARC:UPI000017A3E7
 C/Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
 P/267-520/Domain: protein kinase homology <KIN>
 F/275-283/Region: protein kinase ATP-binding motif

Query Match 7.2%; Score 191; DB 2; Length 544;
 Best Local Similarity 24.6%; Pred. No. 4.4e-05;
 Matches 84; Conservative 60; Mismatches 115; Indels 82; Gaps 16;

4 PEPAPRRAREREREDSESDILE-----SPCGRWKRRQVQGMNPGLOST 55
 Db 230 PPDALTRNTEKQKKRPMSEDE--ILEKLNIVSVGDPKKY--TRFEKIQG--ASGTV 283
 56 FLAMDTEGVEVVNNEHFGDRKAFANHEKIQVFEQVLV--DHPNIVKLHKVLD 112
 Db 284 YTMADVATQGEVAIKQMNLT-----QOPKELINELVMEKNPNIVN---YLD 332
 113 SEACARVIFITEVSSGLKQFLKTKKXHKAMNAKRWCTQILSALSFHACSPPI 172
 Db 333 YLVGDELMVMEYLAGSLTDVYET---CNDQGLIAVCEGLQALEFLH--SNQVI 385
 173 HGNLTSDTIFIQHGLIKIGVWHRIFSNALRPPALPDRLSPIRAREELNHPF-- 230
 Db 386 HRDKSDNILLGMDGSKLDFG--FCAQI-----TPQSKSTVGTVPWMA 431
 231 PEVRYADSTAVDIFSGCALEM-----AVELQNGDTRVTEALARA 276
 Db 432 PVVTRKAYGPKVDISLGMALTEMEGBPPYLNENPLALYIATNG--TPE----- 482
 277 RHLSDPN-----MREPIICCLARDPARPSAHSILFHRV 312
 Db 483 ---LQNEKLSAIFPDILNRCLEMDVYKGSALKELOHPL 520

RESULT 12

T51339
 Mitogen-activated protein kinase kinase (EC 2.7.1.-) 4 [validated] - Arabidopsis thaliana
 N/Alternate names: MAP kinase kinase 4
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 C/Accession: T51339
 R/Ichimura, K.; Mizoguchi, T.; Hayashida, N.; Seki, M.; Shinzaki, K.
 DNA Res. 5, 341-348, 1998
 A/Title: Molecular cloning and characterization of three cDNAs encoding putative mitoge-
 A/Reference number: Z25272; MUID:99156228; PMID:10048483
 A/Accession: T51339
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-366 <ICH>
 A/Cross-references: UNIPROT:O80397; UNIPARC:UPI000009PE2F; EMBL:AB015315; PIDD:BAA28830
 A/Experimental source: strain Columbia
 C/Genetic:
 A/Function:
 C/Description: (EC 2.7.1.-); mitogen-activated protein kinase kinase [validated, MUID:9
 C/Superfamily: kinase-related transforming protein; protein kinase homology
 C/Keywords: phosphotransferase; protein kinase

Query Match 7.1%; Score 190; DB 2; Length 366;
 Best Local Similarity 22.3%; Pred. No. 3.1e-05;
 Matches 81; Conservative 54; Mismatches 122; Indels 106; Gaps 16;

45 NOGNMGLQSTFLAMDTEGVE-----VVMNLEHFGDRKAFAN-- 82
 Db 51 SSGSAPSSGSGASTMTNNSIEAKNYSDLVRGNRIGSGAGCTYKVIHRSSRLVALKVI 110

[illegible]

RESULT 13
G01773
p21-activated protein kinase - human
CISpecies: Homo sapiens (man)
CDate: 21-Dec-1996 #sequence_rev:0n 06-Jun-1997 #text_change 05-Oct-2004
CAccession: G01773
RChernoff, J.
submitted to the EMBL Data Library, April 1995

A:Accession: G01773
A:Structure: Preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-545 <CHE>
A:Cross-references: UNIPROT:Q13153; UNIPARC:UP10000035B53; EMBL:U24152; NID:J9780805; PIR:268-521/Domain: protein kinase homology <KIN>
A:Gene: Pak1

Query Match	7.1%	Score 190;	DB 2;	Length 545;
Best Local Similarity	24.6%;	Pred: No. 5.1e-05;		
Matches 84;	Conservative 60;	Mismatches 115;	Indels 82;	Gaps 16

```

Oy 4 PEPAPRRAREEREDESEDESILAE-----SPCGMOKRREOVONGNMGLOST 55
Db 231 PPDALTTRNTEKQKKPKMSDEE--ILEKLRSLVSGDPRKKY--TRFEKIGG---ASQTV 284
Oy 56 FLANDTEGCEVWVNNELHFGBDKAPAAHEEKIQIVBOLVLV--DHEINYLKHKWIYDT 112
Db 285 YTAMDVAATGAEVAKKNMVL-----OOOPKEELINELIWRKNENIYN-----YIDS 333
Oy 113 SEACARVIFIEVYSSGSILQFLKTKTKKNHAKNARAKMCTOILSLSPFLHACSPPII 172
Db 334 YLVGBELVWVMEYLALAGSLDITVYET-----CMDEGQIAAOCREGLQALYELFH--SNQVI 386
Oy 173 HGNLTSDTIFIQHNGLIKIGSVYHRIFSNALRPTALPDIDRSPIRAEELRLNIHFF-- 230
Db 387 HRDIKSDNILLGMOGSVXLTDTPFG--FCAOI-----TEQSGRSRTMGVTPYVMA 432
Oy 231 PPEYGEVADGTAVDIFSGMCALEM-----AVILEQIOTNDTQVTEBALAR 276
Db 433 PEVTRKRYGKQVDWLSGLIMALEMEBEPPLNENPRALYIATLNG---TBE----- 483
Oy 277 RHSLSDEPN---KREFTLCCILARDPARRPSAHSILFHRVL 312
Db 484 ---LONPEKLSAIFRDFLNRCLMDMVEKRSAGKELQLOFL 521

```

RESULT 14

protein kinase, p21-activated (EC 2.7.1.1-) - human
N/Alternate names: protein Kinase PAK65; S6/H4 kinase
C/Species: Homo sapiens (man)
C/Date: 22-Aug-1996 #sequence revision 06-Sep-1996 #text_change 05-Oct-2004
C/Accession: S58662; S55258; S55304; S58690; A57441
R/Sells: M.; Krause, U.J.; Bagrodia, S.; Ambrose, D.; Bokoch, G.M.; Chernoff, J.
submitted to the EMBL Data Library, April 1995
A/Description: Human p21-activated protein kinases regulate actin organization in mamma
A/Reference number: S58682
A/Accession: S58682
A/Molecule type: DNA
A/Residues: 1-525 <SEU>
A/Cross-references: UNIPROT:Q13177; UNIPARC:UPI0000166649; EMBL:U24153; NID:G780807; PIT
R/Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.
EMBO J. 14, 1970-1978, 1995
A/Title: A novel serine kinase activated by rac1/CDC42Hs-dependent autophosphorylation
A/Reference number: S55258; MUID:952637; PMID:7744004
A/Accession: S55258
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: "METQCKNLHE",31-89,'T','91-149','F',151-224,'T',226-328,'R',330-338,340-525
A/Cross-references: UNIPARC:UPI0000143DCB
A/Accession: S55304
A/Molecule type: protein
A/Residues: 402-418 <MAM>
A/Cross-references: UNIPARC:UPI0000143DCB
R/Martin, G.A.; Bollag, G.; McCormick, F.; Abo, A.
EMBO J. 14, 4385, 1995
A/Reference number: S58690; MUID:96016211; PMID:7556080
A/Content8: erratum
A/Accession: S58690
A/Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-30 <MAF>
A/Cross-references: UNIPARC:UPI000017A3BE
R/Bennet, G.E.; Dennis, P.B.; Masaracchia, R.A.
J. Biol. Chem. 270, 21121-21128, 1995
A/Title: Activation of an S6/H4 kinase (PAK 65) from human placenta by intramolecular ar
A/Reference number: A57441; MUID:95403344; PMID:7673144
A/Accession: A57441
A/Molecule type: protein
A/Residues: 197-216;402,'S',404-409 <BEN>
A/Cross-references: UNIPARC:UPI000017A3BF; UNIPARC:UPI000017A3CO
A/Experimental source: Placenta
C/Keywords: ATP; autophosphorylation; phosphoprotein; phosphotransferase; serine/threonin
F.247-501/domain: protein kinase homology <XIN>
F.255-263/Region: protein kinase ATP-binding motif
F.197,402/binding site: phosphoserine (Ser) (covalent) #status predicted

Query Match 7.0%; Score 188; DB 2; Length 525;
Best Local Similarity 23.6%; Pred. No. 6,6e-05;
Matches 81; Conservative 60; Mismatches 126; Indels 76; Gaps 15;

D 4 PEPA-----RREREREDESEDSOLIE-----SPGRWOKREOVQGN 48
||| ||| : : : : : : : : : : : : : : : : : :
D 201 PVPAVVDSDHDAAGSLDKQKKPKPTDEIMEKRTIVSIGDPKKY-TRYEKIGG- 258
||| ||| : : : : : : : : : : : : : : : : : :
Q 49 MPGLQSTFLMDPBEGEVVMNELHFGRKAFAAHBEKIQTVEQLVL---VDHPNVKL 105
D 259 --ASGTFTALDALGGEVAIKQINL-----OKOPKELIITNEILMKELKNPINV- 308
||| ||| : : : : : : : : : : : : : : : : : :
Q 106 HKYWLDTSEACAVIFPTEVSSGSGLQPLFKTKKHKNKANNAWRKWCTQIISLFHL 165
||| ||| : : : : : : : : : : : : : : : : : :
D 309 ---FLDSYLVGDELFPVMEYLAAGSLTDVTETACMDENAIANV----CRECIALLFFHL 361
||| ||| : : : : : : : : : : : : : : : : : :
Q 166 ACSPPPIIHGULTSDTIPIQHNGLIKIGSVWHRIFSNALRPPTLPDLRSPIAREEELR 225
||| ||| : : : : : : : : : : : : : : : : : :
D 362 A--NOVIHRDIKSDNVALGWEGSVKLTDFG--FCAQI-----TPREGSKSTMV 405
||| ||| : : : : : : : : : : : : : : : : : :
Q 226 NAHFP--PRPYGVADGTANDIFSFGCALFM-----AVLEIQNGDTRVT 269
||| ||| : : : : : : : : : : : : : : : : : :
D 406 GTPEWMAPEVTVTKAAQGPKDWSLGIMALEWEGEPYPYLNNPRLAYLIANNTPELD 465
||| ||| : : : : : : : : : : : : : : : : : :

QY 270 BEAIAARHSLSDPNMREFLICCLARDPARPSAHSILFHRVL 312
 Db 466 NP-----EKLIS-PIFRDPLNRCLEMDVEKRGSAKELLQHPFL 501

RESULT 15

A57597

beta-p21-activated protein kinase - rat

N/Alternate names: beta-PAK

C/Species: Rattus norvegicus (Norway rat)

C/Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 31-Dec-2004

C/Accession: A57597

J./Manager, E.; Chong, C.; Zhao, Z.S.; Leung, T.; Michael, G.; Hall, C.; Lim, L.

J. Biol. Chem. 270, 25070-25078, 1995

A/Title: Molecular cloning of a new member of the p21-Cdc42/Rac-activated kinase (PAK)

A/Reference number: A57597; PMID:96027610; PMID:7559638

A/Accession: A57597

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-544 <MAN>

A/Keywords: ATP

F/266-519/Domain: protein kinase homology <KIN>

F/274-282/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 7.0%; Score 187.5; DB 2; Length 544;

Matches 87; Conservative 57; Mismatches 119; Indels 91; Gaps 17;

QY 2 AAD-BEAPARRA-----REREREDESEDESILRE-----SPCGRWQKRE 42
 Db 214 AAKNKETPPSAENANSTLYRNTDRQKSKMTDEILEKRSIVSGDPKXY-TRFE 272
 QY 43 QVNOGNMPLQSTFLANDTEEGVEVWNNEHFGDRKAFAAHEKIQTFEQVLV---DH 99
 Db 273 KIQGG---ASGTVYTAIDATGQEVAIKQNNL-----QQPKEKELINEILVMRENK 322
 QY 100 PNTVKLHKYLDTSBACARYIFTEYVSSGSLKQFLKTKGNHKAARAKWCTQILS 159
 Db 323 PNTVN---YLDYLVGDELVMVMEYLAGSLTDVYTET-----CMDEGQIAAVCRECLQ 373
 QY 160 ALSFLHACSPPIIHGNLSDTIPIOHNGLIKIGSVWHRIPSNALRPPTALPDDLRSPIRA 219
 Db 374 ALDPLH--SNQVIHRDIDKSDNIIIGMDGSVKLTDFG--FCAOI-----TPEOS 417
 QY 220 EREBELRNLFPP--PPEYGEVADGTAVDIFSGMCALEM-----AVLBIOTN 263
 Db 418 KRSTWGTPTWMAPEVVTTRKAYGKVDIWSLIMATMEVGEPPYLNNENPLRALYLIATN 477
 QY 264 GDRVTVEAIAARHSLSDPN-----MREFLICCLARDPARPSAHSILFHRVL 312
 Db 478 G---TPE-----LQNPRLSAVFRDPLNRCLEMDVDRGSAKELLQHPFL 519

Search completed: January 12, 2006, 11:22:27
 Job time : 20 secs

THIS PAGE BLANK (USPTO)